09/857,841

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/40, 15/82, 5/10, C07K 14/08, A01H 5/00

A1

(11) International Publication Number: WO 96/21018

(43) International Publication Date: 11 July 1996 (11.07.96)

(21) International Application Number: PCT/US95/07234

(22) International Filing Date: 7 June 1995 (07.06.95)

(30) Priority Data:
08/367,789 30 December 1994 (30.12.94) US

(60) Parent Application or Grant
(63) Related by Continuation
US 08/367,789 (CIP)
Filed on 30 December 1994 (30.12.94)

(71) Applicant (*for all designated States except US*): ASGROW SEED COMPANY [US/US]; 2605 East Kilgore Road, Kalamazoo, MI 49002 (US).

(72) Inventors; and
(75) Inventors/Applicants (*for US only*): BOESHORE, Maury, L. [US/US]; 8901 North 24th Street, Kalamazoo, MI 49004 (US). McMASTER, J., Russell [US/US]; 9432 East Main Street, Galesburg, MI 49053 (US). TRICOLI, David, M. [US/US]; 2332 South Rose Street, Kalamazoo, MI 49001 (US). REYNOLDS, John, F. [US/US]; 14815 Trillium Drive, Augusta, MI 49012 (US). CARNEY, Kim, J. [US/US]; 8607 East B Avenue, Richland, MI 49083 (US).

(74) Agent: PERRY, Lawrence, S.; Fitzpatrick, Cella, Harper & Scinto, 277 Park Avenue, New York, NY 10172 (US).

(81) Designated States: AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, JP, KE, KG, KP, KR, KZ, LK, LR, LT, LU, LV, MD, MG, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SI, SK, TJ, TT, UA, US, UZ, VN, ARIPO patent (KE, MW, SD, SZ, UG), European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: PLANTS RESISTANT TO C STRAINS OF CUCUMBER MOSAIC VIRUS

(57) Abstract

Coat protein genes of cucumber mosaic virus strains V27, V33, V34 and A35 (CMV V27, CMV V33, CMV V34, and CMV A35 respectively) are provided.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgystan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic			SE	Sweden
CG	Congo	KR	Republic of Korea	SG	Singapore
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LR	Liberia	SZ	Swaziland
CS	Czechoslovakia	LT	Lithuania	TD	Chad
CZ	Czech Republic	LU	Luxembourg	TG	Togo
DE	Germany	LV	Latvia	TJ	Tajikistan
DK	Denmark	MC	Monaco	TT	Trinidad and Tobago
EE	Estonia	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	UG	Uganda
FI	Finland	ML	Mali	US	United States of America
FR	France	MN	Mongolia	UZ	Uzbekistan
GA	Gabon	MR	Mauritania	VN	Viet Nam

TITLE

PLANTS RESISTANT TO C STRAINS OF CUCUMBER MOSAIC VIRUS

Field of the Invention

This invention relates to coat protein genes derived from cucumber mosaic virus strains V27, V33, V34, and A35 (CMV V27, CMV V33, CMV V34, and CMV A35, respectively). More specifically, the invention relates to the genetic engineering of plants and to a method for conferring viral resistance to a plant using an expression cassette encoding V27, V33, V34, or A35 strains of cucumber mosaic virus.

Background of the Invention

Many agriculturally important crops are susceptible to infection by plant viruses, particularly cucumber mosaic virus, which can seriously damage a crop, reduce its economic value to the grower, and increase its cost to the consumer. Attempts to control or prevent infection of a crop by a plant virus such as cucumber mosaic virus have been made, yet viral pathogens continue to be a significant problem in agriculture.

Scientists have recently developed means to produce virus resistant plants using genetic engineering techniques. Such an approach is advantageous in that the genetic material which provides the protection is incorporated into the genome of the plant itself and can be passed on to its progeny. A host plant is resistant if it possesses the ability to suppress or retard the multiplication of a virus, or the development of pathogenic symptoms. "Resistant" is the opposite of "susceptible," and may be divided into: (1) high, (2) moderate, or (3) low resistance, depending upon its effectiveness. Essentially, a resistant plant shows reduced or no symptom expression, and virus multiplication within it is reduced or negligible. Several different types of host resistance to viruses are recognized. The host may be resistant to: (1) establishment of infection, (2) virus multiplication, or (3) viral movement.

Cucumber mosaic virus (CMV) is a single-stranded (+) RNA plant virus that has a functionally divided genome. The virus genome contains four RNA species designated RNAs 1-4. RNAs 3 and 4 encode the coat protein which is a protein that surrounds the viral RNA and protects the viral RNA from being degraded. Only RNAs 1-3 are required for infectivity because the coat protein, which is encoded by RNA 4, is also encoded by RNA 3.

Several strains of cucumber mosaic virus have been classified using serology, host range, peptide mapping, nucleic acid hybridization, and sequencing analyses. These CMV strains can be divided into two groups, which are designated "WT" (also known as subgroup I) and "S" (also known as subgroup II). The S group consists of at least three members. The WT group is known to contain at least 17 members.

Expression of the coat protein genes from tobacco mosaic virus, alfalfa mosaic virus, cucumber mosaic virus, and potato virus X, among others, in transgenic plants has resulted in plants which are resistant to infection by the respective virus. Heterologous protection can also occur. For example, the expression of coat protein genes from watermelon mosaic virus-2 or zucchini yellow mosaic virus in transgenic tobacco plants has been shown to confer protection against six other potyviruses: bean yellow mosaic virus, potato virus Y, pea mosaic virus, clover yellow vein virus, pepper mottle virus, and tobacco etch virus. However, expression of a preselected coat protein gene does not reliably confer heterologous protection to a plant. For example, transgenic squash plants containing the CMV C coat protein gene, a subgroup I virus, which have been shown to be resistant to the CMV C strain are not protected to the same degree against several highly virulent strains of CMV: CMV V27, CMV V33, CMV V34, and CMV A35 which are all subgroup I viruses.

Thus, a need exists for plants resistant to CMV V27, CMV V33, CMV V34, and CMV A35.

SUMMARY OF THE INVENTION

This invention provides: an isolated and purified DNA molecule that encodes the coat protein for the V27 strain of cucumber mosaic virus (CMV V27), and a chimeric expression cassette comprising this DNA molecule; an isolated and purified DNA molecule that encodes the coat protein for the V33 strain of cucumber mosaic virus (CMV V33), and a chimeric expression cassette comprising this DNA molecule; and an isolated and purified DNA molecule that encodes the coat protein for the V34 strain of cucumber mosaic virus (CMV V34), and a chimeric expression cassette comprising this DNA

molecule; and an isolated and purified DNA molecule that encodes the coat protein for the A35 strain of cucumber mosaic virus (CMV A35), and a chimeric expression cassette comprising the DNA molecule. Another embodiment of the invention is exemplified by the insertion of multiple virus gene expression cassettes into one purified DNA molecule, e.g., a plasmid. Each of these cassettes also includes a promoter which functions in plant cells to cause the production of an RNA molecule, and at least one polyadenylation signal comprising 3' nontranslated DNA which functions in plant cells to cause the termination of transcription and the addition of polyadenylated ribonucleotides to the 3' end of the transcribed mRNA sequences, wherein the promoter is operably linked to the DNA molecule, and the DNA molecule is operably linked to the polyadenylation signal. Preferably, these cassettes include the promoter of the 35S gene of cauliflower mosaic virus and the polyadenylation signal of the cauliflower mosaic virus 35S gene.

Also provided are bacterial cells, and transformed plant cells, containing the chimeric expression cassettes comprising the coat protein genes derived from the CMV V27, CMV V33, CMV V34, or CMV A35 strains, and preferably the 35S promoter of cauliflower mosaic virus and the polyadenylation signal of the cauliflower mosaic virus 35S gene. Plants are also provided, wherein the plants comprise a plurality of transformed cells containing the chimeric coat protein gene expression cassettes derived from the CMV V27, CMV V33, CMV V34, or CMV A35 stains, and preferably the cauliflower mosaic virus 35S promoter and the polyadenylation signal of the cauliflower mosaic virus gene. Transformed plants of this invention include tobacco, beets, corn, cucumber, peppers, potatoes, melons, soybean, squash, and tomatoes. Especially

preferred are members of the *Cucurbitaceae* (e.g., squash and cucumber,) and *Solanaceae* (e.g., peppers and tomatoes) family.

Another aspect of the present invention is a method of preparing a CMV-resistant plant, such as a dicot, comprising: transforming plant cells with a chimeric expression cassette comprising a promoter functional in plant cells operably liked to a DNA molecule that encodes a coat protein as described above; regenerating the plant cells to provide a differentiated plant; and identifying a transformed plant that expresses the CMV coat protein at a level sufficient to render the plant resistant to infection by the specific strains of CMV disclosed herein.

As used herein, with respect to a DNA molecule or "gene," the phrase "isolated and purified" is defined to mean that the molecule is either extracted from its context in the viral genome by chemical means and purified and/or modified to the extent that it can be introduced into the present vectors in the appropriate orientation, i.e., sense or antisense. As used herein, the term "chimeric" refers to the linkage of two or more DNA molecules which are derived from different sources, strains or species (e.g., from bacteria and plants), or the linkage of two or more DNA molecules, which are derived from the same species and which are linked in a way that does not occur in the native genome. As used herein, "expression" is defined to mean transcription or transcription followed by translation of a particular DNA molecule.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1. The nucleotide sequence of the coat protein gene of cucumber mosaic virus V27 [SEQ ID NO:1]. The

deduced amino acid sequence of the encoded open reading frame is shown below the nucleotide sequence [SEQ ID NO:2].

Fig. 2. The nucleotide sequence of the coat protein gene of cucumber mosaic virus V33 [SEQ ID NO:3]. The deduced amino acid sequence of the encoded open reading frame is shown below the nucleotide sequence [SEQ ID NO:4].

Fig. 3. The nucleotide sequence of the coat protein gene of cucumber mosaic virus V34 [SEQ ID NO:5]. The deduced amino acid sequence of the encoded open reading frame is shown below the nucleotide sequence [SEQ ID NO:6].

Fig. 4. The alignment the nucleotide sequences of the coat protein genes from 5 CMV strains [SEQ ID NOS:1, 3, 5, 9, and 10]. Ccp and Cmvwl[SEQ ID NO:9 and 10] are described in Quemada et al. (J. Gen. Virol., 70, 1065 (1989)). Alignments were obtained with the use of the UWGCG Pileup program. The dots represent either the lack of sequence information at the 5' end of the coat protein gene or gaps in homology in sequences relative to others in the alignment. The positions of primers RMM351 and RMM352 are shown [SEQ ID NOS:7 and 8].

Fig. 5. The alignment of the amino acid sequences deduced from the nucleotide sequences of CMV strains V27, V33, V34, CMV-C (shown in Fig. 4 [SEQ ID NO:1, 3, 5, 9 and 10]) and CMV strain Cmvq3 (Quemada et al., J. Gen. Virol., 70, 1065 (1989)) [SEQ ID NO:2, 4, 6, 11 and 12]. Alignments were performed by the UWGCG Pileup program. Differences among the "C" type viruses are underlined and highlighted with asterisks. The dots represent gaps in homology in sequences relative to others in the alignment.

Fig. 6. (A) Assembly of CMV strain V27 coat protein expression cassette. PCR products of CMV V27 were installed into pCRII and subsequently inserted into pUC18cpexpress by routine methods. The bolded lines and arrows which are a part of the circle represent CaMV 35S sequences. (B) Insertion of a CMV V27 coat protein expression cassette BamHI fragment into the BglII site of pEPG204 and pEPG205 to produce pEPG239 and pEPG240, respectively. (C) Restriction map of pEPG239. This binary plasmid includes the coat protein expression cassettes for PRV (melon, long), CMV V27, ZYMV, and WMVII. For further information on PRV coat protein genes, refer to Applicants' Assignees copending Patent Application Serial No. 08/366,881 entitled "Papaya Ringspot Virus Coat Protein Gene" filed on December 30, 1994, incorporated by reference herein. For further information on ZYMV and WMVII coat protein genes, refer to Applicants' Assignees copending Patent Application Serial No. 08/232,846 filed on April 25, 1994 entitled "Potyvirus Coat Protein Genes and Plants Transformed Therewith", incorporated by reference herein. (D) Restriction map of pEPG240. This binary plasmid includes the coat protein expression cassettes for PRV (melon, short), CMV V27, ZYMV, and WMVII.

Fig. 7. (A) Assembly of CMV strain V33 coat protein expression cassette. PCR products of CMV V33 were installed into pUC1318cpexpress by routine methods. (B) Insertion of a CMV V33 coat protein expression cassette BamHI fragment into the BglII site of pEPG204 and pEPG205 to produce pEPG196 and pEPG197, respectively. (C) Restriction map of pEPG196. This binary plasmid includes the coat protein expression cassettes for PRV (melon, long), CMV V33, ZYMV, and WMVII. Arrows indicate CaMV 35S promoter fragments. (D) Restriction map of pEPG197. This binary plasmid

includes the coat protein expression cassettes for PRV (melon, short), CMV V33, ZYMV, and WMVII.

Fig. 8. The nucleotide sequence of the coat protein gene of cucumber mosaic virus A35 [SEQ ID NO:14]. The deduced amino acid sequence of the encoded open reading frame is shown below the nucleotide sequence [SEQ ID NO:15].

Fig. 9. The alignment of the amino acid sequences deduced from the nucleotide sequences of the six CMV strains shown in Fig. 10 [SEQ ID NO:2, 4, 6, 11, 12 and 15]. Differences among the "C" type viruses are underlined and highlighted with asterisks. The dots represent gaps in homology in sequences relative to others in the alignment.

Fig. 10. The alignment the nucleotide sequences of the coat protein genes from 6 CMV strains [SEQ ID NOS:1, 3, 5, 9, 10 and 14]. The dots represent either the lack of sequence information at the 5' end of the coat protein gene or gaps in homology in sequences relative to others in the alignment.

DETAILED DESCRIPTION OF THE INVENTION

Cucumber mosaic virus (CMV) is a single-stranded (+) RNA plant virus that has a functionally divided genome. The virus genome contains four RNA species designated RNAs 1-4; 3389 nucleotides (nt), 3035 nt, 2193 nt, and 1027 nt, respectively (Peden et al., Virol., 53, 487 (1973); Gould et al., Eur. J. Biochem., 126, 217 (1982); Rezaian et al., Eur. J. Biochem., 143, 227 (1984); Rezaian et al., Eur. J. Biochem. 150, 331

(1985)). Only RNAs 1-3 are required for infectivity (Peden et al., Virol., 53, 487 (1973)) because the coat protein, which is encoded by RNA 4, is also encoded by RNA 3. Translations of CMV RNAs yield a 95 kD polypeptide from RNA 1, a 94 kD polypeptide from RNA 2 (Gordon et al., Virol., 123, 284 (1983)), and two polypeptides from RNA 3: its 5' end encodes a 35 kD polypeptide, and its 3' end encodes a 24.5 kD polypeptide (Gould et al., Eur. J. Biochem., 126, 217 (1982)). The 24.5 kD polypeptide is identical to that encoded by RNA 4 and is the coat protein.

Several strains of cucumber mosaic virus have been classified using serology, host range, peptide mapping, nucleic acid hybridization, and sequencing. These CMV strains can be divided into two groups, which are designated "WT" (also known as subgroup I) and "S" (also known as subgroup II). CMV subgroup I includes CMV-C, CMV-V27, CMV-V33, CMV-V34, CMV-M, CMV-O, CMV-Y, and CMV-A35 while subgroup II includes CMV-Q, CMV-WL, and CMV-LS (Zaitlin et al., Virol., 201, 200 (1994)). Protection against a strain in one group does not necessarily provide protection against all strains in that group. For example, transgenic squash plants protected with coat protein genes from the CMV strain C are not protected against the CMV strains V27, V33, V34, or A35. In addition, Zaitlin et al. (Virol., 201, 200 (1994)) report that tobacco plants transgenic for a CMV-FNY replicase gene show protection against challenge from subgroup I strains but show no protection against challenge from subgroup II challenges. Thus, the present invention is directed to providing plants with resistance to CMV strains V27, V33, V34, and/or A35.

To practice the present invention, a viral gene must be isolated from the viral genome and inserted into a

vector. Thus, the present invention provides isolated and purified DNA molecules that encode the coat proteins of the V27, V33, or V34 strains of CMV. As used herein, a DNA molecule that encodes a coat protein gene includes nucleotides of the coding strand, also referred to as the "sense" strand, as well as nucleotides of the noncoding strand, complementary strand, also referred to as the "antisense" strand, either alone or in their base-paired configuration. Thus, a DNA molecule that encodes the coat protein of the V27 strain of CMV, for example, includes the DNA molecule having the nucleotide sequence of Figure 1 [SEQ ID NO:1], a DNA molecule complementary to the nucleotide sequence of Figure 1 [SEQ ID NO:1], as well as a DNA molecule which also encodes a CMV coat protein and its complement which hybridizes with a CMV V27-specific DNA probe in hybridization buffer with 6XSSC, 5X Denhardt's reagent, 0.5% SDS and 100 μg/ml denatured, fragmented salmon sperm DNA and remains bound when washed at 68°C in 0.1XSSC and 0.5% SDS (Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd ed. (1989)). Moreover, the DNA molecules of the present invention can include non-CMV coat protein nucleotides that do not interfere with expression of the CMV coat protein gene. Preferably, the isolated and purified DNA molecules of the present invention comprise a single coding region for the coat protein. Thus, preferably the DNA molecules of the present invention are those "consisting essentially of" DNA that encodes the coat protein.

These CMV genes are used to produce the coat proteins, which are believed to confer resistance to viruses. Another molecular strategy to provide virus resistance in transgenic plants is based on antisense RNA. As is well known, a cell manufactures protein by transcribing the DNA of the gene encoding that protein to produce

RNA, which is then processed to messenger RNA (mRNA) (e.g., by the removal of introns) and finally translated by ribosomes into protein. This process may be inhibited in the cell by the presense of antisense RNA. The term antisense RNA means an RNA sequence which is complementary to a sequence of bases in the mRNA in question in the sense that each base (or the majority of bases) in the antisense sequence (read in the 3' to 5' sense) is capable of pairing with the corresponding base (G with C, A with U) in the mRNA sequence read in the 5' to 3' sense. It is believed that this inhibition takes place by formation of a complex between the two complementary strands of RNA, thus preventing the formation of protein. How this works is uncertain: the complex may interfere with further transcription, processing, transport or translation, or degrade the mRNA, or have more than one of these effects. This antisense RNA may be produced in the cell by transformation of the cell with an appropriate DNA construct arranged to transcribe the non-template strand (as opposed to the template strand) of the relevant gene (or of a DNA sequence showing substantial homology therewith).

The use of antisense RNA to downregulate the expression of specific plant genes is well known. Reduction of gene expression has led to a change in the phenotype of the plant: either at the level of gross visible phenotypic difference, e.g., lack of anthocyanin production in flower petals of petunia leading to colorless instead of colored petals (van der Krol et al., Nature, 333:866-869 (1988)); or at a more subtle biochemical level, e.g., change in the amount of polygalacturonase and reduction in depolymerization of pectin during tomato fruit ripening (Smith et al., Nature, 334:724-726 (1988)).

Another more recently described method of inhibiting gene expression in transgenic plants is the use of sense RNA transcribed from an exogenous template to downregulate the expression of specific plant genes (Jorgensen, Keystone Symposium "Improved Crop and Plant Products through Biotechnology", Abstract X1-022 (1994)). Thus, both antisense and sense RNA have been proven to be useful in achieving downregulation of gene expression in plants, which are encompassed by the present invention.

The CMV coat protein gene does not contain the signals necessary for its expression once transferred and integrated into a plant genome. Accordingly, a vector must be constructed to provide the regulatory sequences such that they will be functional upon inserting a desired gene. When the expression vector/insert construct is assembled, it is used to transform plant cells which are then used to regenerate plants. These transgenic plants carry the viral gene in the expression vector/insert construct. The gene is expressed in the plant and increased resistance to viral infection is conferred thereby.

Several different methods exist to isolate a viral gene. To do so, one having ordinary skill in the art can use information about the genomic organization of cucumoviruses to locate and isolate the coat protein gene. The coat protein gene is located near the 3' end of RNA 3. Using methods well known in the art, a quantity of virus is grown and harvested. The viral RNA is then separated by gel electrophoresis. A cDNA library is created using the viral RNA, by methods known to the art. The viral RNA is incubated with primers that hybridize to the viral RNA and reverse transcriptase, and a complementary DNA molecule is produced. A DNA complement of the complementary DNA

molecule is produced and that sequence represents a DNA copy (cDNA) of the original viral RNA molecule. The DNA complement can be produced in a manner that results in a single double stranded cDNA or polymerase chain reactions can be used to amplify the DNA encoding the cDNA with the use of oligomer primers specific for viral sequences. These primers can include novel restriction sites used in subsequent cloning steps. Thus, a double stranded DNA molecule is generated which contains the sequence information of the viral RNA. These DNA molecules can be cloned in *E. coli* plasmid vectors after the additions of restriction enzyme linker molecules by DNA ligase. The various fragments are inserted into cloning vectors, such as well-characterized plasmids, which are then used to transform *E. coli* and create a cDNA library.

CMV coat protein genes from previously isolated strains can be used as hybridization probes to screen the cDNA library to determine if any of the transformed bacteria contain DNA fragments with sequences coding for a CMV coat protein. Alternatively, plasmids which harbor CMV coat protein sequences can be determined by restriction enzyme digestion of plasmids in bacterial transformants. The cDNA inserts in any bacterial colonies which contain this region can be sequenced. The coat protein gene is present in its entirety in colonies which have sequences that extend 5′ to the sequence which encodes the ATG start codon and sequences that extend 3′ of the stop codon.

Alternatively, cDNA fragments can be inserted in the sense orientation into expression vectors. Antibodies against the coat protein can be used to screen the cDNA expression library and the gene can be isolated from colonies which express the protein.

In the present invention, the DNA molecules encoding the coat protein (CP) genes of the cucumber mosaic virus strains V27, V33, V34, and A35 have been determined and the genes have been inserted into expression cassettes. These expression cassettes can be individually placed into a vector that can be transmitted into plants, preferably a binary vector. Alternatively, two or more of the CMV CP genes can each be present in an expression cassette which can be placed into the same binary vector, or any of the CMV CP expression cassettes of the present invention can be placed into a binary vector with one or more viral gene expression cassettes. The expression vectors contain the necessary genetic regulatory sequences for expression of an inserted gene. The coat protein gene is inserted such that those regulatory sequences are functional and the genes can be expressed when incorporated into a plant genome. For example, vectors of the present invention can contain combinations of expression cassettes that include DNA from a heterologous CMV coat protein gene (i.e., one from another CMV isolate), papaya ringspot virus coat protein gene, a zucchini yellow mosaic virus coat protein gene, and a watermelon mosaic virus-2 coat protein gene.

Moreover, when combinations of viral gene expression cassettes are placed in the same binary plasmid, and that multigene cassette containing plasmid transformed into a plant, the viral genes all preferably exhibit substantially the same degrees of efficacy when present in transgenic plants. For example, if one examines numerous transgenic lines containing two different intact viral gene cassettes, the transgenic line will be immune to infection by both viruses. Similarly, if a line exhibits a delay in symptom development to one virus, it will also exhibit a delay in symptom

development to the second virus. Finally, if a line is susceptible to one of the viruses it will be susceptible to the other. This phenomenon is unexpected. If there were not a correlation between the efficacy of each gene in these multiple gene constructs this approach as a tool in plant breeding would probably be prohibitively difficult to use. Even with single gene constructs, one must test numerous transgenic plant lines to find one that displays the appropriate level of efficacy. The probability of finding a line with useful levels of expression can range from 10-50% (depending on the species involved). For further information refer to Applicants' assignees copending Patent Application Serial No. 08/367,788 entitled "Transgenic Plants Expressing DNA Constructs Containing a Plurality of Genes to Impart Virus Resistance" filed on December 30, 1994, incorporated by reference herein.

In order to express the viral gene, the necessary genetic regulatory sequences must be provided. In the present invention, the coat protein genes are inserted into vectors which contain cloning sites for insertion 3' of the initiation codon and 5' of the poly(A) signal. The promoter is 5' of the initiation codon such that when genes are inserted at the cloning site, a functional unit is formed in which the inserted genes are expressed under the control of the various genetic regulatory sequences.

The segment of DNA referred to as the promoter is responsible for the regulation of the transcription of DNA into mRNA. A number of promoters which function in plant cells are known in the art and can be employed in the practice of the present invention. These promoters can be obtained from a variety of sources such as plants or plant viruses, and can include, but are not

limited to, promoters isolated from the caulimovirus group such as the cauliflower mosaic virus 35S promoter (CaMV 35S), the enhanced cauliflower mosaic virus 35S promoter (enh CaMV35S), the figwort mosaic virus full-length transcript promoter (FMV35S), and the promoter isolated from the chlorophyll a/b binding protein. Other useful promoters include promoters which are capable of expressing the cucumovirus proteins in an inducible manner or in a tissue-specific manner in certain cell types in which the infection is known to occur. For example, the inducible promoters from phenylalanine ammonia lyase, chalcone synthase, hydroxyproline rich glycoprotein, extensin, pathogenesis-related proteins (e.g. PR-1a), and wound-inducible protease inhibitor from potato may be useful.

Preferred promoters for use in the present CP-containing cassettes include the constitutive promoters from CaMV, the Ti genes nopaline synthase (Bevan et al., Nucleic Acids Res. II, 369 (1983)) and octopine synthase (Depicker et al., J. Mol. Appl. Genet., 1, 561 (1982)), and the bean storage protein gene phaseolin. The poly(A) addition signals from these genes are also suitable for use in the present cassettes. The particular promoter selected is preferably capable of causing sufficient expression of the DNA coding sequences to which it is operably linked, to result in the production of amounts of the proteins or RNA effective to provide viral resistance, but not so much as to be detrimental to the cell in which they are expressed. The promoters selected should be capable of functioning in tissues including, but not limited to, epidermal, vascular, and mesophyll tissues. The actual choice of the promoter is not critical, as long as it has sufficient transcriptional activity to accomplish the expression of the preselected proteins or their

respective RNAs and subsequent conferral of viral resistance to the plants.

The nontranslated leader sequence can be derived from any suitable source and can be specifically modified to increase the translation of the mRNA. The 5' nontranslated region can be obtained from the promoter selected to express the gene, an unrelated promoter, the native leader sequence of the gene or coding region to be expressed, viral RNAs, suitable eucaryotic genes, or a synthetic gene sequence. The present invention is not limited to the constructs presented in the following examples.

The termination region or 3' nontranslated region which is employed is one which will cause the termination of transcription and the addition of polyadenylated ribonucleotides to the 3' end of the transcribed mRNA sequence. The termination region can be native with the promoter region, native with the gene, or can be derived from another source, and preferably include a terminator and a sequence coding for polyadenylation. Suitable 3' nontranslated regions of the chimeric plant gene include but are not limited to: (1) the 3' transcribed, nontranslated regions containing the polyadenylation signal of *Agrobacterium* tumor-inducing (Ti) plasmid genes, such as the nopaline synthase (NOS) gene; and (2) plant genes like the soybean 7S storage protein genes.

Preferably, the expression cassettes of the present invention are engineered to contain a constitutive promoter 5' to its translation initiation codon (ATG) and a poly(A) addition signal (AATAAA) 3' to its translation termination codon. Several promoters which function in plants are available, however, the preferred promoter is the 35S constitutive promoters

from cauliflower mosaic virus (CaMV). The poly (A) signal can be obtained from the CaMV 35S gene or from any number of well characterized plant genes, i.e., nopaline synthase, octopine synthase, and the bean storage protein gene phaseolin. The constructions are similar to that used for the expression of the CMV C coat protein in PCT Patent Application PCT/US88/04321, published on June 29, 1989 as WO 89/05858, claiming the benefit of U.S. SN 135,591, filed December 21, 1987, entitled "Cucumber Mosaic Virus Coat Protein Gene", and the CMV WL coat protein in PCT Patent Application PCT/US89/03288, published on March 8, 1990 as WO 90/02185, claiming the benefit of U.S. SN 234,404, filed August 19, 1988, entitled "Cucumber Mosaic Virus Coat Protein Gene."

Selectable marker genes can be incorporated into the present expression cassettes and used to select for those cells or plants which have become transformed. The marker gene employed may express resistance to an antibiotic, such as kanamycin, gentamycin, G418, hygromycin, streptomycin, spectinomycin, tetracyline, chloramphenicol, and the like. Other markers could be employed in addition to or in the alternative, such as, for example, a gene coding for herbicide tolerance such as tolerance to glyphosate, sulfonylurea, phosphinothricin, or bromoxynil. Additional means of selection could include resistance to methotrexate, heavy metals, complementation providing prototrophy to an auxotrophic host, and the like.

The particular marker employed will be one which will allow for the selection of transformed cells as opposed to those cells which are not transformed. Depending on the number of different host species one or more markers can be employed, where different conditions of selection would be useful to select the different host,

and would be known to those of skill in the art. A screenable marker such as the β-glucuronidase gene can be used in place of, or with, a selectable marker. Cells transformed with this gene can be identified by the production of a blue product on treatment with 5-bromo-4-chloro-3-indoyl-β-D-glucuronide (X-Gluc).

In developing the present expression construct, i.e., expression cassette, the various components of the expression construct such as the DNA molecules, linkers, or fragments thereof will normally be inserted into a convenient cloning vector, such as a plasmid or phage, which is capable of replication in a bacterial host, such as *E. coli*. Numerous cloning vectors exist that have been described in the literature. After each cloning, the cloning vector can be isolated and subjected to further manipulation, such as restriction, insertion of new fragments, ligation, deletion, resection, insertion, *in vitro* mutagenesis, addition of polylinker fragments, and the like, in order to provide a vector which will meet a particular need.

For *Agrobacterium*-mediated transformation, the expression cassette will be included in a vector, and flanked by fragments of the *Agrobacterium* Ti or Ri plasmid, representing the right and, optionally the left, borders of the Ti or Ri plasmid transferred DNA (T-DNA). This facilitates integration of the present chimeric DNA sequences into the genome of the host plant cell. This vector will also contain sequences that facilitate replication of the plasmid in *Agrobacterium* cells, as well as in *E. coli* cells.

All DNA manipulations are typically carried out in *E. coli* cells, and the final plasmid bearing the cucumovirus expression cassette is moved into *Agrobacterium* cells by direct DNA transformation,

conjugation, and the like. These *Agrobacterium* cells will contain a second plasmid, also derived from Ti or Ri plasmids. This second plasmid will carry all the vir genes required for transfer of the foreign DNA into plant cells. Suitable plant transformation cloning vectors include those derived from a Ti plasmid of *Agrobacterium tumefaciens*, as generally disclosed in Glassman et al. (U.S. Pat. No. 5,258,300), or *Agrobacterium rhizogenes*.

A variety of techniques are available for the introduction of the genetic material into or transformation of the plant cell host. However, the particular manner of introduction of the plant vector into the host is not critical to the practice of the present invention, and any method which provides for efficient transformation can be employed. In addition to transformation using plant transformation vectors derived from the tumor-inducing (Ti) or root-inducing (Ri) plasmids of *Agrobacterium*, alternative methods could be used to insert the DNA constructs of the present invention into plant cells. Such methods may include, for example, the use of liposomes, electroporation (Fromm et al., Proc. Natl. Acad. Sci. USA, 82, 824 (1984)), chemicals that increase the free uptake of DNA (Paszkowski et al., EMBO J., 3, 2717 (1984)), DNA delivery via microprojectile bombardment (Klein et al., Nature, 327, 70 (1987)), microinjection (Crossway et al., Mol. Gen. Genet., 202, 179 (1985)), and transformation using viruses or pollen.

The choice of plant tissue source or cultured plant cells for transformation will depend on the nature of the host plant and the transformation protocol. Useful tissue sources include callus, suspension culture cells, protoplasts, leaf segments, stem segments, tassels, pollen, embryos, hypocotyls, tuber segments,

meristematic regions, and the like. The tissue source is regenerable, in that it will retain the ability to regenerate whole, fertile plants following transformation.

The transformation is carried out under conditions directed to the plant tissue of choice. The plant cells or tissue are exposed to the DNA carrying the present viral gene expression cassette(s) for an effective period of time. This can range from a less-than-one-second pulse of electricity for electroporation, to a two-to-three day co-cultivation in the presence of plasmid-bearing *Agrobacterium* cells. Buffers and media used will also vary with the plant tissue source and transformation protocol. Many transformation protocols employ a feeder layer of suspended culture cells (tobacco or Black Mexican Sweet Corn, for example) on the surface of solid media plates, separated by a sterile filter paper disk from the plant cells or tissues being transformed.

Following treatment with DNA, the plant cells or tissue may be cultivated for varying lengths of time prior to selection, or may be immediately exposed to a selective agent such as those described hereinabove. Protocols involving exposure to *Agrobacterium* will also include an agent inhibitory to the growth of the *Agrobacterium* cells. Commonly used compounds are antibiotics such as cefotaxime and carbenicillin. The media used in the selection may be formulated to maintain transformed callus or suspension culture cells in an undifferentiated state, or to allow production of shoots from callus, leaf or stem segments, tuber disks, and the like.

Cells or callus observed to be growing in the presence of normally inhibitory concentrations of the selective

agents are presumed to be transformed and may be subcultured several additional times on the same medium to remove nonresistant sections. The cells or calli can then be assayed for the presence of the viral gene cassette, or can be subjected to known plant regeneration protocols. In protocols involving the direct production of shoots, those shoots appearing on the selective media are presumed to be transformed and can be excised and rooted, either on selective medium suitable for the production of roots, or by simply dipping the excised shoot in a root-inducing compound and directly planting it in vermiculite.

In order to produce transgenic plants exhibiting viral resistance, the viral genes must be taken up into the plant cell and stably integrated within the plant genome. Plant cells and tissues selected for their resistance to an inhibitory agent are presumed to have acquired the selectable marker gene encoding this resistance during the transformation treatment. Since the marker gene is commonly linked to the viral genes, it can be assumed that the viral genes have similarly been acquired. Southern blot hybridization analysis using a probe specific to the viral genes can then be used to confirm that the foreign genes have been taken up and integrated into the genome of the plant cell. This technique may also give some indication of the number of copies of the gene that have been incorporated. Successful transcription of the foreign gene into mRNA can likewise be assayed using Northern blot hybridization analysis of total cellular RNA and/or cellular RNA that has been enriched in a polyadenylated region. mRNA molecules encompassed within the scope of the invention are those which contain viral specific sequences derived from the viral genes present in the transformed vector which are of the same polarity as that of the viral genomic RNA such

that they are capable of base pairing with viral specific RNA of the opposite polarity to that of viral genomic RNA under conditions described in Chapter 7 of Sambrook et al. (1989). Moreover, mRNA molecules encompassed within the scope of the invention are those which contain viral specific sequences derived from the viral genes present in the transformed vector which are of the opposite polarity as that of the viral genomic RNA such that they are capable of base pairing with viral genomic RNA under conditions described in Chapter 7 in Sambrook et al. (1989).

The presence of a viral gene can also be detected by immunological assays, such as the double-antibody sandwich assays described by Namba et al., Gene, 107, 181 (1991) as modified by Clark et al., J. Gen. Virol., 34, 475 (1979). See also, Namba et al., Phytopathology, 82, 940 (1992). Cucumovirus resistance can also be assayed via infectivity studies as generally disclosed by Namba et al., ibid., wherein plants are scored as symptomatic when any inoculated leaf shows veinclearing, mosaic or necrotic symptoms.

Seed from plants regenerated from tissue culture is grown in the field and self-pollinated to generate true breeding plants. The progeny from these plants become true breeding lines which are evaluated for viral resistance in the field under a range of environmental conditions. The commercial value of viral-resistant plants is greatest if many different hybrid combinations with resistance are available for sale. The farmer typically grows more than one kind of hybrid based on such differences as maturity, color or other agronomic traits. Additionally, hybrids adapted to one part of a country are not adapted to another part because of differences in such traits as maturity, disease and insect tolerance. Because of this, it is

necessary to breed viral resistance into a large number of parental lines so that many hybrid combinations can be produced.

The invention will be further described by reference to the following detailed examples. Enzymes were obtained from commercial sources and were used according to the vendor's recommendations or other variations known in the art. Other reagents, buffers, etc., were obtained from commercial sources, such as Sigma Chemical Co., St. Louis, MO, unless otherwise specified.

Most of the recombinant DNA methods employed in practicing the present invention are standard procedures, well known to those skilled in the art, and described in detail in, for example, in European Patent Application Publication Number 223,452, published November 29, 1986, which is incorporated herein by reference. General references containing such standard techniques include the following: R. Wu, ed., Methods in Enzymology, Vol. 68 (1979); J.H. Miller, Experiments in Molecular Genetics (1972); J. Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd ed. (1989); and D.M. Glover, ed., DNA Cloning Vol. II (1982).

Figures 6 and 7 are presented to illustrate the constructions of this invention.

Example I.

A. Isolation of CMV RNAs

Zucchini squash plants (20-day old) were inoculated with CMV strains V27, V33, or V34; after 7-10 days, infected leaves were harvested and CMV virus particles

were isolated. The procedure used was based on protocols from Lot et al., <u>Annals of Phytopathology</u>, <u>4</u>, 25 (1972), Francki et al., <u>CMI/AAB Descriptions of Plant Viruses</u>, (July, 1979), and Habili and Francki, <u>Virology</u>, <u>57</u>, 292 (1974). Approximately 100 g of fresh leaves were extracted in an equal volume (w/v) of 0.5 M Na-citrate (pH 6.5) containing 5 mM EDTA and 100 mL of chloroform. After centrifugation of the extract at 12,000 x g for 10 minutes, polyethyleneglycol ("PEG", Sigma Chemical Co. PEG-8000, average molecular weight, Research Grade) was added to the supernatant to a final concentration of 10% and the suspension was stirred for 30-40 minutes at 0-4°C. This suspension was centrifuged at 12,000 x g for 10 minutes, and the pellet was resuspended in 40-50 mL of 5 mM Na-borate buffer (pH 9.0) containing 0.5 M EDTA. TRITON X-100 was then added to the the virus particle suspension to a final concentration of 2% and stirred on ice for 30 minutes. This suspension was then centrifuged at 19,000 x g for 15 minutes, and the supernatant was collected and subsequently centrifuged at 105,000 x g for 2 hours. The virus pellet was collected and resuspended in about 2 mL of 5 mM Na-borate buffer (pH 9.0) containing 0.5 mM EDTA. The resuspended virus preparation was applied onto a step sucrose gradient consisting of 5 layers: 5%, 10%, 15%, 20%, and 25% sucrose dissolved in 2.0 mM Na-phosphate buffer (pH 7.5). Gradients were centrifuged at 37,000 rpm in a Sorvall TH641 swinging bucket rotor for 45 minutes. After centrifugation, the virus band was harvested, the virus preparation was dialyzed against Na-borate buffer, and LiCl was added (2M final concentration) to lyse the virions and to precipitate viral RNA. CMV RNA was dissolved and reprecipitated with ethanol and dissolved in water. By agarose gel electrophoresis, the expected four RNA species were observed.

B. Cloning CMV Coat Protein Genes

(a) CMV V27

The first cDNA strand of CMV V27 was synthesized with the use of Perkin-Elmer RT-PCR kit reagents and the primer RMM352 (shown in Figure 4, [SEQ ID NO:8]); immediately in the same reaction tube, a polymerase chain reaction (PCR) was carried out with the use of oligonucleotide primers RMM351 and RMM352 (shown in Figure 4, [SEQ ID NOS:7 and 8] following the manufacturer's protocol. The ATG translation start is included in the NcoI site present in primer RMM351. Individual PCR product molecules were cloned using the TA Cloning™ kit (Invitrogen Corp., San Diego, CA) into pCRII (included in the TA Cloning™ kit as a linearized plasmid with single 3' dT overhangs at the ends of the molecule). Three clones were isolated for further study: CMVV27TA21, CMVV27TA23, and CMVV27TA26. With the use of a kit (Sequenase 2 purchased from USB, Cleveland, Ohio), the CMV V27 insert in clone CMVV27TA21 was sequenced.

CMMV27 was compared to 11 different CMV isolates: Cmvbaul, Cmvq3, Cmvwl, Cmvtrk7, Cmvfc, Cmvil7f, Cmvc, Cmvpr50, Cmvv27, Cmvp6, Cmvo, Cmvm, and Cmvy. CMVV27 coat protein is similar to CMV-Y in that it contains a serine at position 29 while other strains have an alanine at this position. However, CMV-Y contains a leucine at position 18 while CMVV27 contains a proline at position 18. In addition, CMVV27 has a methionine at position 206, no other CMV-C group viruses have a methionine at this position (Baulcombe, D., "Mutational analysis of CMV RNA3: Effects on RNA3 accumulation, RNA4 synthesis and plant infection." Unpublished Direct Submission. Submitted (19-JUN-1992) David Baulcombe, The Sainsbury Laboratory, Norwich Research Park, Colney Lane, Norwich, NR2 7UH, United Kingdom; Hayakawa et al., Gene, 71, 107 (1988);

Hayakawa et al., J. Gen. Virol. 70, 499 (1989); Owen et al., J. Gen. Virol., 71, 2243 (1990); Pappu et al., "The nucleotide and the deduced amino acid sequences of coat protein genes of three Puerto Rican isolates of cucumber mosaic virus." Unpublished (1992). This sequence is included in the GeneBank sequence data base; Salanki et al., "Complete nucleotide sequence of RNA 3 from cucumber mosaic virus strain Trk 7." Unpublished (1993). This sequence is included in the GeneBank data base; Shintaku, J. Gen. Virol. 72, 2587 (1991)).

(b) CMV V33

CMV V33 was purified and viral RNA extracted from a virion preparation as described above; subsequently single stranded cDNA was synthesized using Perkin-Elmer RT-PCR kit reagents and oligomer primer RMM352 [SEQ ID NO:8]. The coat protein gene of strain V33 was amplified using PCR as described above for V27 with the use of oligomer primers RMM351 and RMM352 (Figure 4, [SEQ ID NOS:7 and 8, respectively]). The V33 CP gene PCR product was digested with NcoI and directly cloned into the expression cassette cpexpress installed into pUC1318 (see Kay and McPherson, Nucleic Acid Research, 15, 2779 (1987) for pUC1318; Slightom, Gene 100, 251 (1991) for cpexpress; pUC1318cpexpress is the cpexpress described in Slightom, however it is installed into the HindIII site of the modified pUC plasmid pUC1318 described in detail in Kay and McPherson), rather than into the intermediate vector pCRII. By colony hybridization with a CMV coat protein probe, a number of clones were identified for further analysis: V33ce1, V33ce2, V33ce7, and V33ce9. The CMV V33 insert in clone V33ce7 was sequenced with the use of a kit (Sequenase 2 purchased from USB, Cleveland, Ohio).

CMMV33 was compared to 11 different CMV isolates: Cmvbaul, Cmvq3, Cmvwl, Cmvtrk7, Cmvfc, Cmvil7f, Cmvc, Cmvpr50, Cmvv27, Cmvp6, Cmvo, Cmvm, and Cmvy. CMVV33 has a serine at position 67 while all other CMV strains compared included a proline at this position. At position 196, both CMVV33 and CMV-Y have a valine residue; all other members of the CMV-C group contains isoleucine at this position. However, at position 184, CMVV33 has an alanine residue while CMV-Y has a threonine residue. Therefore, CMVV33 coat protein is unique (Baulcombe, D., "Mutational analysis of CMV RNA3: Effects on RNA3 accumulation, RNA4 synthesis and plant infection." Unpublished Direct Submission. Submitted (19-JUN-1992) David Baulcombe, The Sainsbury Laboratory, Norwich Research Park, Colney Lane, Norwich, NR2 7UH, United Kingdom; Hayakawa et al., Gene, 71, 107 (1988); Hayakawa et al., J. Gen. Virol. 70, 499 (1989); Owen et al., J. Gen. Virol., 71, 2243 (1990); Pappu et al., "The nucleotide and the deduced amino acid sequences of coat protein genes of three Puerto Rican isolates of cucumber mosaic virus." Unpublished (1992). This sequence is included in the GeneBank sequence data base; Salanki et al., "Complete nucleotide sequence of RNA 3 from cucumber mosaic virus strain Trk 7." Unpublished (1993). This sequence is included in the GeneBank data base; Shintaku, J. Gen. Virol. 72, 2587 (1991)).

(c) CMV V34

CMV V34 RNA was prepared as described above. Subsequently, the first cDNA strand was synthesized using CMV V34 template in a reaction that included the following: approximately 2 μg CMV V34 RNA, 1 x buffer for Superscript Reverse Transcriptase (supplied by BRL-GIBCO, Grand Island, NY), 2 mM dNTPs, oligomer primer RMM352 (37.5 μg/mL, SEQ ID NO:8), 1.5 μL RNasin, and 1 μL Superscript Reverse Transcriptase (BRL-GIBCO) in a

20-μL reaction. After this reaction was allowed to proceed for 30 minutes, an aliquot of the first strand reaction was used as a template in a polymerase chain reaction to amplify the CMV V34 coat protein gene. The CMV V34 coat protein gene PCR product was cloned into the pCRII vector included in the TA Cloning™ Kit supplied by Invitrogen Corp. Two clones were isolated for further study: TA17V34 and TA112V34. The CMV V34 insert of clone TA17V34 was sequenced with the use of a kit (Sequenase 2 purchased from USB, Cleveland, Ohio). Comparative sequence analysis of the CMVV34 coat protein gene with other CMV coat protein genes (Cmvbaul, Cmvq3, Cmvwl, Cmvtrk7, Cmvfc, Cmvil7f, Cmvc, Cmvpr50, Cmvv27, Cmvp6, Cmvo, Cmvm, and Cmvy) showed that the CMVV34 coat protein gene is unique (Baulcombe,D. Mutational analysis of CMV RNA3: Effects on RNA3 accumulation, RNA4 synthesis and plant infection. Unpublished Direct Submission. Submitted (19-JUN-1992) David Baulcombe, The Sainsbury Laboratory, Norwich Research Park, Colney Lane, Norwich, NR2 7UH, United Kingdom; Hayakawa et al., Gene, 71, 107 (1988); Hayakawa et al., J. Gen. Virol. 70, 499 (1989); Owen et al., J. Gen. Virol., 71, 2243 (1990); Pappu et al., (1992) The nucleotide and the deduced amino acid sequences of coat protein genes of three Puerto Rican isolates of cucumber mosaic virus. Unpublished. This sequence is included in the GeneBank sequence data base; Salanki et al., Complete nucleotide sequence of RNA 3 from cucumber mosaic virus strain Trk 7. Unpublished (1993) This sequence is included in the GeneBank data base; Shintaku, J. Gen. Virol. 72, 2587 (1991)).

C. Engineering CMV Coat Protein Genes

(a) CMV V27

The NcoI fragment in CMVV27TA21 that harbors CMVV27 CP coding sequences was excised from CMVV27TA21 and

inserted into the plant expression cassette cpexpress in pUC18 to give CMVV27TA21ce42. The resulting expression cassette was isolated as a partial HindIII fragment and inserted into the binary vector pGA482G [The parent binary plasmid was pGA482, constructed by An (Plant Physiol., 81, 86 (1986)). This binary vector contains the T-DNA border sequences from pTiT37, the selectable marker gene Nos-NPT II (which contains the plant-expressible nopaline gene promoter fused to the bacterial NPT II gene obtained from Tn5), a multiple cloning region, and the cohesive ends of phage lambda (An, Plant Physiol., 81, 86 (1986))] to yield pEPG191 and pEPG192. Subsequently, a PRV coat protein expression cassette was installed to obtain a binary vector that included both CMV V27 CP and PRV CP expression cassettes.

Alternatively, the CMV V27 CP NcoI fragment obtained from CMV V27TA21 was installed into pUC1318cp express (see Kay et al., Nucleic Acid Research, 15, 2779 (1987) for pUC1318; Slightom, Gene 100, 251 (1991) for cpexpress; pUC1318cpexpress is the cpexpress described in Slightom, however it is installed into the HindIII site of the modified pUC plasmid pUC 1318 described in detail in Kay et al.) to give the plasmid CMVV27TA21CE13 (similar to CMVV27TA21ce42). The plasmid pUC1318 provided additional sites (e.g., BamHI and Xbal) with which the cassette could be inserted into the binary vector pGA482G Subsequently, the bacteria-derived gentamicin-(3)-N-acetyl-transferase gene (Allmansberger et al., Mol. Gen. Genet., 198, 514 (1985)) was installed into a SalI site outside of the T-DNA region, adjacent to the left border (B_L)). The BamHI fragment harboring the CMV strain V27 CP expression cassette was isolated and inserted into the BglII site of the binary plasmid pEPG205 (PRV34/Z72/WMBN22) to give pEPG240 (CMVV27/PRV34/Z72/WMBN22). The BamHI fragment was also

installed into the BglII site of the binary plasmid pEPG204 (PRV16/Z72/WMBN22) to yield pEPG239 (CMVV2716/PRV16/Z72/WMBN22) (Table 1). For further information on PRV coat protein genes, refer to Applicants' assignees copending Patent Application Serial No. 08/366,881 entitled "Papaya Ringspot Virus Coat Protein Gene" filed on December 30, 1994, incorporated by reference herein. For further information on ZYMV and WMVII coat protein genes, refer to Applicants' assignees copending Patent Application Serial No. 08/232,846 filed on April 25, 1994 entitled "Potyvirus Coat Protein Genes and Plants Transformed Therewith", incorporated by reference herein.

Table 1

Binary	Parental Plasmid	Site	CMVcp Cassette	pEPG#
pGA482G	pGA482G	HindIII	CMVV27cpexpress	191 or 192
pPRBN	pEPG204 (P16sZW)	BglII	CMVV27cpexpress	239
pPRBN	pEPG204 (P16sZW)	BglII	CMVV27cpexpress	240
pPRBN	pEPG106 (ZW)	HindIII	CMVV27cpexpress	243
pGA482G	pGA482G	HindIII	CMVV33ce7	198
pPRBN	pEPG106 (ZW)	HindIII	CMVV33ce7	244
pPRBN	pEPG204 (P16sZW)	BglII	CMVV27ce7	196
pPRBN	pEPG205 (P34sZW)	BglII	CMVV27ce7	197
pGA482G	pGA482G	HindIII	17V34cpexp113	190

(b) <u>CMV V33</u>

Subsequently, both HindIII and BamHI fragments were excised from clone V33ce7; these fragments carried the complete expression cassette for CMV V33 CP gene. The BamHI fragment (V33 CP expression cassette) was inserted into the BglII site of pEPG204 (PRV16/ZY72/WMBN22) to obtain pEPG196. The BamHI fragment was also inserted into the BglII site of pEPG205 (PRV34/ZY72/WMBN22) to obtain pEPG197 (V3329/PRV34/ZY72/WMBN22). The HindIII fragment

harboring the V33 CP cassette was installed into pGA482G to obtain pEPG198 (Table 1).

(c) <u>CMV V34</u>

An NcoI fragment excised from clone TA17V34 was installed into the NcoI site of pUC1318 cpexpress. A resulting plasmid that includes the CMV V34 coding NcoI fragment inserted in the sense orientation is 17V34/cpexp113. A partial HindIII fragment from the plasmid 17V34/cpexp113 was isolated and installed into pGA482G to yield pEPG190 (Table 1).

(d) <u>Agrobacterium Strains</u>

The binary plasmids described here, such as pPRBN (for further information on these plasmids, refer to Applicants' Assignees copending Patent Application Serial No. 08/366,991 entitled "Transgenic Plants Expressing DNA Constructs Containing a Plurality of Genes to Impart Virus Resistance" filed on December 30, 1994, incorporated by reference herein) or their derivatives, can be transferred into Agrobacterium strains A208, C58, LBA4404, C58Z707, A4RS, A4RS(pRi278b), Mog301 and others. Strains A208, C58, LBA4404, and A4RS are available from ATCC, 12301 Parklawn Drive, Rockville, Maryland. A4RS (pRi278b) was obtained from Dr. F. Casse-Delbart, C.N.R.A., Route de Saint Cyr, F78000, Versailles, France. C58Z707 was obtained from Dr. A.G. Hepburn, University of Illinois, Urbana, Illinois. Mog301 was obtained from Mogen NV, Leiden, Netherlands.

D. <u>Transfer of CMV Coat Protein Genes to Tobacco</u>

In order to test whether the CMV CP gene constructs described herein confer protection against CMV challenge with homologous strains, some of the binary plasmids listed above (e.g., pEPG197, pEPG198, pEPG239,

and pEPG240) have been used to insert these novel CMV coat protein genes into *Nicotiana tobacum*. Agrobacterium-mediated transfer of the plant expressible CMV coat protein genes described herein was done using the methods described in PCT published application WO 89/05859, entitled "Agrobacterium Mediated Transformation of Germinating Plant Seeds".

Five R_1 progeny lines of *Nicotiana t.* transformed with the binary plasmid pEPG239 and five R_1 progeny lines of *Nicotiana t.* transformed with the binary plasmid pEPG240 have been obtained. These binary plasmids include the coat protein gene of CMV strain V27. The ten R_0 parental plants of these lines were assayed for NPTII protein expression by ELISA. They each expressed NPTII protein by ELISA. Furthermore, these ten lines were assayed for both the NPTII and CMV V27 coat protein genes by PCR analysis. PCR analysis detected both genes in all ten R_0 plants.

The binary plasmid pEPG198 was used to obtain 11 R_0transgenic *Nicotiana t.* plants. By PCR analysis, the CMV V33 CP gene was detected in nine of the eleven R_0 plants tested.

<u>Cloning and engineering CMV A35 CP Gene</u>

20-day-old zucchini squash plants in the greenhouse were inoculated with CMV strain A35; after 7-10 days infected leaves were harvested. Total RNA was isolated from these infected plants by the use of Tri-Reagent and the instructions provided with the reagent (Molecular Research Center, inc., Cincinnati, OH). Single-stranded cDNA was synthesized using total RNA template. The reaction included 1 X first Strand cDNA Synthesis Buffer (GIBCO-BRL), 1mM dNTP's (Pharmacia), 2 uL oligonucleotide primer RMM352 (150ug/mL), 2 uL

RNasin (Promega), and 1uL RTase SuperscriptII (GIBCO-BRL) in a 20uL reaction volume. The CMV A35 coat protein gene was PCR amplified with the use of CMV coat protein-specific primers RMM351 and 352 [SEQ ID NOS:7 and 8]. The PCR included 3uL of the cDNA synthesis reaction described above, 8 uL of each primer RMM351 and RMM352 (150 ug/uL stock), 5uL 10X reaction buffer, 4uL dNTP's (10mM), 1.5 uL $MgCl_2$ (50mM), and 0.5 uL Taq polymerase (BRL-GIBCO). PCR conditions were carried out as follows: 93° 45 sec, 50° 45 sec, then 72° 180 sec for 30 cycles, then 72° for 5 min, then hold at 4°. PCR products were visualized by agarose gel electrophoresis and subsequently cloned.

PCR product molecules were cloned into the pCRII vector supplied with the TA cloning kit (Invitrogen Corp.) Four clones were identified and restriction mapped, however, none were sequenced for further analysis.

Alternatively, an aliquot of the CMV A35 PCR product was digested with NcoI and ligated it into the NcoI site of pUC19B2 cp express to give the plasmid CMV A35cpexp33. The cost protein insert of this plasmid was sequenced with the use of the Sequenase II Kit supplied by USBiochemical (Figure 8). Sequence analysis reveals that CMV A35 coat protein sequence differs form the coast protein sequences of CMV C, V27, V33, V34, and WL (Figures 9 and 10). For example, A35 differs from other CMV C strains at amino acid position #26 (Figure 9). Examination of the nucleotide sequence comparisons differs from other CMV coat protein genes characterized (Figure 10).

A BamHI/BI1II fragment was excised from A35cpexp33 and installed into the unique BglII site of pGA482G. The plasmid pUC19B2cpexp provides a BamHI site at the 5' end of the cpexp cassette and a BglII site at the 3'

end of the expression cassette. Upon insertion into a BglII site, the unique BglII site of the binary plasmid pGA482 is maintained for subsequent insertions of gene cassettes. Binary plasmids that include the CMV A35 expression cassette are being transformed into various *Agrobacterium* strains (eg., C58Z707, Mog301, and LBA4404). These *Agrobacterium* strains are used to transform plants to impart resistance to CMV CARNA5.

All publications, patents and patent documents are incorporated by reference herein, as though individually incorporated by reference. The invention has been described with reference to various specific and preferred embodiments and techniques. However, it should be understood that many variations and modifications may be made while remaining within the spirit and scope of the invention.

WHAT IS CLAIMED IS:

1. In isolated and purified DNA molecule consisting essentially of DNA encoding the coat protein of the V27 strain of cucumber mosaic virus.

2. The isolated and purified DNA molecule of claim 1 wherein the DNA molecule has the nucleotide sequence shown in Figure 1 [SEQ ID NO:1].

3. A vector comprising a chimeric expression cassette comprising the DNA molecule of claim 1, a promoter and a polyadenylation signal, wherein the promoter is operably linked to the DNA molecule, and the DNA molecule is operably linked to the polyadenylation signal.

4. The vector of claim 3 wherein the promoter is the cauliflower mosaic virus 35S promoter.

5. The vector of claim 4 wherein the polyadenylation signal is the polyadenylation signal of the cauliflower mosaic 35S gene.

6. A bacterial cell comprising the vector of claim 3.

7. The bacterial cell of claim 6 wherein the bacterial cell is selected from the group consisting of an *Agrobacterium tumefaciens* cell and an *Agrobacterium rhizogenes* cell.

8. A transformed plant cell transformed with the vector of claim 3.

9. The transformed plant cell of claim 8 wherein the promoter is cauliflower mosaic virus 35S promoter and

the polyadenylation signal is the polyadenylation signal of the cauliflower mosaic 35S gene.

10. A plant selected from the family *Cucurbitaceae* comprising a plurality of the transformed cells of claim 8.

11. A plant selected from the family *Solanaceae* comprising a plurality of the transformed cells of claim 8.

12. An isolated and purified DNA molecule consisting essentially of DNA encoding the coat protein of the V33 strain of cucumber mosaic virus.

13. The isolated and purified DNA molecule of claim 12 wherein the DNA
molecule has the nucleotide sequence shown in Figure 2 [SEQ ID NO:3].

14. A vector comprising a chimeric expression cassette comprising the DNA molecule of claim 12, a promoter and a polyadenylation signal, wherein the promoter is operably linked to the DNA molecule, and the DNA molecule is operably linked to the polyadenylation signal.

15. The vector of claim 14 wherein the promoter is the cauliflower mosaic virus 35S promoter.

16. The vector of claim 15 wherein the polyadenylation signal is the polyadenylation signal of the cauliflower mosaic 35S gene.

17. A bacterial cell comprising the vector of claim 14.

18. The bacterial cell of claim 17 wherein the bacterial cell is selected from the group consisting of an *Agrobacterium tumefaciens* cell and an *Agrobacterium rhizogenes* cell.

19. A transformed plant cell transformed with the vector of claim 14.

20. The transformed plant cell of claim 19 wherein the promoter is cauliflower mosaic virus 35S promoter and the polyadenylation signal is the polyadenylation signal of the cauliflower mosaic 35S gene.

21. A plant selected from the family *Cucurbitaceae* comprising a plurality of the transformed cells of claim 19.

22. A plant selected from the family *Solanaceae* comprising a plurality of the transformed cells of claim 19.

23. An isolated and purified DNA molecule consisting essentially of DNA encoding the coat protein of the V34 strain of cucumber mosaic virus.

24. The isolated and purified DNA molecule of claim 23 wherein the DNA

molecule has the nucleotide sequence shown in Figure 3 [SEQ ID NO:5].

25. A vector comprising a chimeric expression cassette comprising the DNA molecule of claim 24, a promoter and a polyadenylation signal, wherein the promoter is operably linked to the DNA molecule, and the DNA molecule is operably linked to the polyadenylation signal.

26. The vector of claim 25 wherein the promoter is cauliflower mosaic virus 35S promoter.

27. The vector of claim 26 wherein the polyadenylation signal is the polyadenylation signal of the cauliflower mosaic 35S gene.

28. A bacterial cell comprising the vector of claim 23.

29. The bacterial cell of claim 28 wherein said bacterial cell is selected from the group consisting of an *Agrobacterium tumefaciens* cell and an *Agrobacterium rhizogenes* cell.

30. A transformed plant cell transformed with the vector of claim 25.

31. The transformed plant cell of claim 30 wherein the promoter is cauliflower mosaic virus 35S promoter and the polyadenylation signal is the polyadenylation signal of the cauliflower mosaic 35S gene.

32. A plant selected from the family *Cucurbitaceae* comprising a plurality of the transformed cells of claim 30.

33. A plant selected from the family *Solanaceae* comprising a plurality of the transformed cells of claim 30.

34. A method of preparing a cucumber mosaic viral resistant plant comprising:

(a) transforming plant cells with a chimeric expression cassette comprising a promoter functional in plant cells operably linked to a DNA molecule that encodes a coat protein; wherein the DNA molecule is

derived from a cucumber mosaic virus strain selected from the group consisting of V27, V33, and V34;

(b) regenerating the plant cells to provide a differentiated plant; and

(c) identifying a transformed plant that expresses the cucumber mosaic virus coat protein at a level sufficient to render the plant resistant to infection by the cucumber mosaic virus strain.

35. The method of claim 34 wherein the plant is a dicot.

36. The method of claim 35 wherein the dicot is selected from the family *Cucurbitaceae*.

37. The method of claim 35 wherein the dicot is selected from the family *Solanaceae*.

38. A vector comprising a chimeric expression cassette comprising the DNA molecule of claim 1 and at least one chimeric expression cassette comprising a heterologous CMV coat protein gene, a papaya ringspot virus coat protein gene, a zucchini yellow mosaic virus coat protein gene, or a watermelon mosaic virus-2 coat protein gene, wherein each expression cassette comprises a promoter and a polyadenylation signal, wherein the promoter is operably linked to the DNA molecule, and the DNA molecule is operably linked to the polyadenylation signal

39. A bacterial cell comprising the vector of claim 38.

40. A transformed plant cell transformed with the vector of claim 38.

41. The transformed plant cell of claim 40 wherein the promoter is cauliflower mosaic virus 35S promoter and the polyadenylation signal is the polyadenylation signal of the cauliflower mosaic 35S gene.

42. A vector comprising a chimeric expression cassette comprising the DNA molecule of claim 12 and at least one chimeric expression cassette comprising a heterologous CMV coat protein gene, a papaya ringspot virus coat protein gene, a zucchini yellow mosaic virus coat protein gene, or a watermelon mosaic virus-2 coat protein gene, wherein each expression cassette comprises a promoter and a polyadenylation signal, wherein the promoter is operably linked to the DNA molecule, and the DNA molecule is operably linked to the polyadenylation signal

43. A bacterial cell comprising the vector of claim 42.

44. A transformed plant cell transformed with the vector of claim 42.

45. The transformed plant cell of claim 44 wherein the promoter is cauliflower mosaic virus 35S promoter and the polyadenylation signal is the polyadenylation signal of the cauliflower mosaic 35S gene.

46. A vector comprising a chimeric expression cassette comprising the DNA molecule of claim 23 and at least one chimeric expression cassette comprising a heterologous CMV coat protein gene, a papaya ringspot virus coat protein gene, a zucchini yellow mosaic virus coat protein gene, or a watermelon mosaic virus-2 coat protein gene, wherein each expression cassette comprises a promoter and a polyadenylation signal, wherein the promoter is operably linked to the DNA molecule, and the DNA molecule is operably linked to the polyadenylation signal

47. A bacterial cell comprising the vector of claim 46.

48. A transformed plant cell transformed with the vector of claim 46.

49. The transformed plant cell of claim 48 wherein the promoter is cauliflower mosaic virus 35S promoter and the polyadenylation signal is the polyadenylation signal of the cauliflower mosaic 35S gene.

FIG. 1A

```
  1  CCATGGACAAATCTGAATCAACCAGTGCTGGTCGTAACCGTCGGCGTCGTCCGCGTCGTG   60
      MetAspLysSerGluSerThrSerAlaGlyArgAsnArgArgArgArgProArgArgG
      M  D  K  S  E  S  T  S  A  G  R  N  R  R  R  R  P  R  R  G

 61  GTTCCCGCTCCGCCTCCTCCTCCTCGGATGCTAACTTTAGAGTCTTGTCGCAGCAGCTTT  120
     lySerArgSerAlaSerSerSerSerAspAlaAsnPheArgValLeuSerGlnGlnLeuS
      S  R  S  A  S  S  S  S  D  A  N  F  R  V  L  S  Q  Q  L  S

121  CGCGACTTAACAAGACGTTAGCAGCTGGTCGTCCAACTATTAACCACCCAACCTTTGTAG  180
     erArgLeuAsnLysThrLeuAlaAlaGlyArgProThrIleAsnHisProThrPheValG
      R  L  N  K  T  L  A  A  G  R  P  T  I  N  H  P  T  F  V  G

181  GGAGTGAACGCTGTAAACCTGGGTACACGTTCACATCTATTACCCTAAAGCCACCAAAAA  240
     lySerGluArgCysLysProGlyTyrThrPheThrSerIleThrLeuLysProProLysI
      S  E  R  C  K  P  G  Y  T  F  T  S  I  T  L  K  P  P  K  I

241  TAGACCGTGGGTCTTATTACGGTAAAAGGTTGTTATTACCTGATTCAGTCACGGAATATG  300
     leAspArgGlySerTyrTyrGlyLysArgLeuLeuLeuProAspSerValThrGluTyrA
      D  R  G  S  Y  Y  G  K  R  L  L  L  P  D  S  V  T  E  Y  D

301  ATAAGAAGCTTGTTTCGCGCATTCAAATTCGAGTTAATCCTTTGCCGAAATTTGATTCTA  360
     spLysLysLeuValSerArgIleGlnIleArgValAsnProLeuProLysPheAspSerT
      K  K  L  V  S  R  I  Q  I  R  V  N  P  L  P  K  F  D  S  T
```

FIG. 1B

```
361  CCGTGTGGGTAACAGTCCGTAAAGTTCCTGCCTCCTCGGACTTATCCGTTGCCGCCATCT  420
     hrValTrpValThrValArgLysValProAlaSerSerAspLeuSerValAlaAlaIleS
       V  W  V  T  V  R  K  V  P  A  S  S  D  L  S  V  A  A  I  S

421  CTGCTATGTTCGCGGACGGAGCCTCACCGGTACTGGTTTATCAGTATGCTGCATCTGGAG  480
     erAlaMetPheAlaAspGlyAlaSerProValLeuValTyrGlnTyrAlaAlaSerGlyV
       A  M  F  A  D  G  A  S  P  V  L  V  Y  Q  Y  A  A  S  G  V

481  TCCAAGCTAACAACAAATTGTTGTATGATCTTTCGGCGATGCGCGCTGATATAGGTGACA  540
     alGlnAlaAsnAsnLysLeuLeuTyrAspLeuSerAlaMetArgAlaAspIleGlyAspM
       Q  A  N  N  K  L  L  Y  D  L  S  A  M  R  A  D  I  G  D  M

541  TGAGAAAGTACGCCGTCCTCGTGTATTCAAAAGACGATGCGCTCGAGACGGACGAGCTAG  600
     etArgLysTyrAlaValLeuValTyrSerLysAspAspAlaLeuGluThrAspGluLeuV
       R  K  Y  A  V  L  V  Y  S  K  D  D  A  L  E  T  D  E  L  V

601  TACTTCATGTTGACATCGAGCACCAACGTATTCCCACGTCTGGGATGCTCCCAGTCTGAT  660
     alLeuHisValAspIleGluHisGlnArgIleProThrSerGlyMetLeuProValEnd
       L  H  V  D  I  E  H  Q  R  I  P  T  S  G  M  L  P  V  *

661  TCCGTGTTCCCAGAACCCTCCCTCCGATTTCTGTGGCGGGAGCTGAGTTGGCAGTTCTGC  720
721  TATAAACTGTCTGAAGTCACTAAACGTTTCACGGTGAACGGGTTGTCCATGG  772
```

FIG. 2A

```
  1  CCATGGACAAATCTGAATCAACCAGTGCTGGTCGTAACCGTCGACGTCGTCCGCGTCGTG  60
      MetAspLysSerGluSerThrSerAlaGlyArgAsnArgArgArgArgProArgArgG
      M  D  K  S  E  S  T  S  A  G  R  N  R  R  R  R  P  R  R  G

 61  GTTCCCGCTCCGCCCCCTCCTCCGCGGATGCCAACTTTAGAGTCTTGTCGCAGCAGCTTT  120
     lySerArgSerAlaProSerSerAlaAspAlaAsnPheArgValLeuSerGlnGlnLeuS
      S  R  S  A  P  S  S  A  D  A  N  F  R  V  L  S  Q  Q  L  S

121  CGCGACTTAATAAGACGTTGTCAGCTGGTCGTCCAACTATTAACCACCCAACCTTTGTAG  180
     erArgLeuAsnLysThrLeuSerAlaGlyArgProThrIleAsnHisProThrPheValG
      R  L  N  K  T  L  S  A  G  R  P  T  I  N  H  P  T  F  V  G

181  GGAGTGAGCGTTGTAAATCTGGGTACACGTTCACATCTATTACCCTAAAGCCGCCGAAAA  240
     lySerGluArgCysLysSerGlyTyrThrPheThrSerIleThrLeuLysProProLysI
      S  E  R  C  K  S  G  Y  T  F  T  S  I  T  L  K  P  P  K  I

241  TAGACCGTGGGTCTTATTATGGTAAAAGGTTGTTATTACCTGATTCAGTCACAGAATATG  300
     leAspArgGlySerTyrTyrGlyLysArgLeuLeuLeuProAspSerValThrGluTyrA
      D  R  G  S  Y  Y  G  K  R  L  L  L  P  D  S  V  T  E  Y  D

301  ATAAGAAACTTGTTTCGCGCATTCAAATTCGAGTTAATCCCTTGCCGAAATTTGATTCTA  360
     spLysLysLeuValSerArgIleGlnIleArgValAsnProLeuProLysPheAspSerT
      K  K  L  V  S  R  I  Q  I  R  V  N  P  L  P  K  F  D  S  T

361  CCGTGTGGGTGACAGTCCGTAAAGTTCCTGCCTCCTCGGACTTATCCGTTGCCGCCATCT  420
     hrValTrpValThrValArgLysValProAlaSerSerAspLeuSerValAlaAlaIleS
      V  W  V  T  V  R  K  V  P  A  S  S  D  L  S  V  A  A  I  S
```

FIG. 2B

```
421  CTGCTATGTTTGCGGACGGAGCCTCACCGGTACTGGTTTATCAGTACGCTGCATCTGGAG  480
     erAlaMetPheAlaAspGlyAlaSerProValLeuValTyrGlnTyrAlaAlaSerGlyV
      A  M  F  A  D  G  A  S  P  V  L  V  Y  Q  Y  A  A  S  G  V

481  TCCAAGCTAACAACAAATTGTTGTATGATCTTTCGGCGATGCGCGCTGATATAGGCGACA  540
     alGlnAlaAsnAsnLysLeuLeuTyrAspLeuSerAlaMetArgAlaAspIleGlyAspM
      Q  A  N  N  K  L  L  Y  D  L  S  A  M  R  A  D  I  G  D  M

541  TGAGAAAGTACGCCGTCCTCGTGTATTCAAAAGACGATGCACTCGAGACGGACGAGCTAG  600
     etArgLysTyrAlaValLeuValTyrSerLysAspAspAlaLeuGluThrAspGluLeuV
      R  K  Y  A  V  L  V  Y  S  K  D  D  A  L  E  T  D  E  L  V

601  TACTTCATGTTGACGTCGAGCACCAACGCATTCCCACGTCTGGGGTGCTCCCAGTATAAT  660
     alLeuHisValAspValGluHisGlnArgIleProThrSerGlyValLeuProValEnd
      L  H  V  D  V  E  H  Q  R  I  P  T  S  G  V  L  P  V  *
661  TCTGTGCTTTCCAGAACCCTCCCTCCGATTTCTGTGGCGGGAGCTGAGTTGGCAGTTCTG  720
721     CTGTAAACTGTCTGAAGTCACTAAACGTTTTACGGTGAACGGGTTGTCCATGG  773
```

FIG. 3

```
1   CCATGGACAAATCTGAATCAACCAGTGCTGGTCGTAACCGTCGACGTCGTCCGCGTCGTGGTTCCCGCTCCGCTTCCTCCTCTTCGGATGCTAACTTTAG   100
     MetAspLysSerGluSerThrSerAlaGlyArgAsnArgArgArgArgProArgArgGlySerArgSerAlaSerSerSerSerAspAlaAsnPheAr
     M  D  K  S  E  S  T  S  A  G  R  N  R  R  R  R  P  R  R  G  S  R  S  A  S  S  S  S  D  A  N  F  R

101 AGTCTTGTCGCAGCAGCTTTCGCGACTTAACAAGACGTTAGCAGCTGGTCGTCCAACTATTAACCACCCAACCTTTGTAGGGAGTGAACGCTGTAGACCT   200
    gValLeuSerGlnGlnLeuSerArgLeuAsnLysThrLeuAlaAlaGlyArgProThrIleAsnHisProThrPheValGlySerGluArgCysArgPro
      V  L  S  Q  Q  L  S  R  L  N  K  T  L  A  A  G  R  P  T  I  N  H  P  T  F  V  G  S  E  R  C  R  P

201 GGGTACACGTTCACATCTATTACCCTAAAGCCACCAAAAATAGACCGCGGGTCTTACTACGGTAAAAGGTTGTTACTACCTGATTCAGTCACGGAATATG   300
    GlyTyrThrPheThrSerIleThrLeuLysProProLysIleAspArgGlySerTyrTyrGlyLysArgLeuLeuLeuProAspSerValThrGluTyrA
    G  Y  T  F  T  S  I  T  L  K  P  P  K  I  D  R  G  S  Y  Y  G  K  R  L  L  L  P  D  S  V  T  E  Y  D

301 ATAAGAAGCTTGTTTCGCGCATTCAAATTCGAGTTAATCCTTTGCCGAAATTTGATTCTACCGTGTGGGTGACAGTTCGTAAAGTTCCTGCCTCCTCGGA   400
    spLysLysLeuValSerArgIleGlnIleArgValAsnProLeuProLysPheAspSerThrValTrpValThrValArgLysValProAlaSerSerAs
      K  K  L  V  S  R  I  Q  I  R  V  N  P  L  P  K  F  D  S  T  V  W  V  T  V  R  K  V  P  A  S  S  D

401 CTTATCCGTTGCCGCCATCTCTGCTATGTTCGCGGACGGAGCCTCACCGGTACTGGTTTATCAGTATGCTGCATCTGGAGTTCAAGCTAACAACAAATTG   500
    pLeuSerValAlaAlaIleSerAlaMetPheAlaAspGlyAlaSerProValLeuValTyrGlnTyrAlaAlaSerGlyValGlnAlaAsnAsnLysLeu
      L  S  V  A  A  I  S  A  M  F  A  D  G  A  S  P  V  L  V  Y  Q  Y  A  A  S  G  V  Q  A  N  N  K  L

501 TTGTATGATCTTTCGGCGATGCGCGCTGATATAGGTGACATGAGAAAGTACGCCGTCCTCGTGTATTCAAAAGACGATGCACTCGAGACGGACGAGCTAG   600
    LeuTyrAspLeuSerAlaMetArgAlaAspIleGlyAspMetArgLysTyrAlaValLeuValTyrSerLysAspAspAlaLeuGluThrAspGluLeuV
    L  Y  D  L  S  A  M  R  A  D  I  G  D  M  R  K  Y  A  V  L  V  Y  S  K  D  D  A  L  E  T  D  E  L  V

601 TACTTCATGTTGACATCGAGCACCAACGCATTCCCACGTCTGGGGTGCTCCCAGTTTGATTCCGTGTTCCAGAACCCTCCCTCCGATTTCTGTGGCGGGA   700
    alLeuHisValAspIleGluHisGlnArgIleProThrSerGlyValLeuProValEnd
      L  H  V  D  I  E  H  Q  R  I  P  T  S  G  V  L  P  V  *

701 GCTGAGTTGGCAGTTCTGCTATAAACTGTCTGAAGTCACTAAACGTTTTACGGTGAACGGGTTGTCCATGG   771
```

FIG. 4A

```
    RMM351              NcoI
5' CGTAGAATTCAGTCG AGCCATGGAC 3'                                420
 V27cp  .......... ..CCATGGAC AAATCTGAAT CAACCAGTGC TGGTCGTAAC CGTCGGCGTC
 V33cp  .......... ..CCATGGAC AAATCTGAAT CAACCAGTGC TGGTCGTAAC CGTCGACGTC
Cmvv34  .......... ..CCATGGAC AAATCTGAAT CAACCAGTGC TGGTCGTAAC CGTCGACGTC
   Ccp  AATTGAGTCG AGTCATGGAC AAATCTGAAT CAACCAGTGC TGGTCGTAAC CATCGACGTC
Cmvwl   GTCTTAGTGT GCCTATGGAC AAATCTGGAT CTCCCAATGC TAGTAGAACC TCCCGGCGTC

        421                                                     480
 V27cp  GTCCGCGTCG TGGTTCCCGC TCCGCCTCCT CCTCCTCGGA TGCTAACTTT AGAGTCTTGT
 V33cp  GTCCGCGTCG TGGTTCCCGC TCCGCCCCCT CCTCCGCGGA TGCCAACTTT AGAGTCTTGT
Cmvv34  GTCCGCGTCG TGGTTCCCGC TCCGCTTCCT CCTCTTCGGA TGCTAACTTT AGAGTCTTGT
   Ccp  GTCCGCGTCG TGGTTCCCGC TCCGCCCCCT CCTCCGCGGA TGCTAACTTT AGAGTCTTGT
Cmvwl   GTCGCCCGCG TAGAGGTTCT CGGTCCGCTT CTGGTGCGGA TGCAGGGTTG CGTGCTTTGA

        481                                                     540
 V27cp  CGCAGCAGCT TTCGCGACTT AACAAGACGT TAGCAGCTGG TCGTCCAACT ATTAACCACC
 V33cp  CGCAGCAGCT TTCGCGACTT AATAAGACGT TGTCAGCTGG TCGTCCAACT ATTAACCACC
Cmvv34  CGCAGCAGCT TTCGCGACTT AACAAGACGT TAGCAGCTGG TCGTCCAACT ATTAACCACC
   Ccp  CGCAGCAGCT TTCGCGACTT AATAAGACGT TAGCAGCTGG TCGTCCAACT ATTAACCACC
Cmvwl   CTCAGCAGAT GCTGAAACTC AATAGAACCC TCGCCATTGG TCGTCCCACT CTTAACCACC

        541                                                     600
 V27cp  CAACCTTTGT AGGGAGTGAA CGCTGTAAAC CTGGGTACAC GTTCACATCT ATTACCCTAA
 V33cp  CAACCTTTGT AGGGAGTGAG CGTTGTAAAT CTGGGTACAC GTTCACATCT ATTACCCTAA
Cmvv34  CAACCTTTGT AGGGAGTGAA CGCTGTAGAC CTGGGTACAC GTTCACATCT ATTACCCTAA
   Ccp  CAACCTTTGT AGGGAGTGAA CGCTGTAGAC CTGGGTACAC GTTCACATCT ATTACCCTAA
Cmvwl   CAACCTTCGT GGGTAGTGAA AGCTGTAAAC CCGGTTACAC TTTCACATCT ATTACCCTGA
```

FIG. 4B

```
         601                                                        660
V27cp    AGCCACCAAA AATAGACCGT GGGTCTTATT ACGGTAAAAG GTTGTTATTA CCTGATTCAG
V33cp    AGCCGCCGAA AATAGACCGT GGGTCTTATT ATGGTAAAAG GTTGTTATTA CCTGATTCAG
Cmvv34   AGCCACCAAA AATAGACCGC GGGTCTTACT ACGGTAAAAG GTTGTTACTA CCTGATTCAG
Ccp      AGCCACCAAA AATAGACCGT GAGTCTTATT ACGGTAAAAG GTTGTTACTA CCTGATTCAG
Cmvwl    AACCGCCTGA AATTGAGAAA GGTTCATATT TTGGTAGAAG GTTGTCTTTG CCAGATTCAG

         661                                                        720
V27cp    TCACGGAATA TGATAAGAAG CTTGTTTCGC GCATTCAAAT TCGAGTTAAT CCTTTGCCGA
V33 :p   TCACAGAATA TGATAAGAAA CTTGTTTCGC GCATTCAAAT TCGAGTTAAT CCCTTGCCGA
Cmvv34   TCACGGAATA TGATAAGAAG CTTGTTTCGC GCATTCAAAT TCGAGTTAAT CCTTTGCCGA
Ccp      TCACGGAATA TGATAAGAAG CTTGTTTCGC GCATTCAAAT TCGAGTTAAT CCTTTGCCGA
Cmvwl    TCACGGACTA TGATAAGAAG CTTGTTTCGC GCATTCAAAT CAGGGTTAAT CCTTTGCCGA

         721                                                        780
V27cp    AATTTGATTC TACCGTGTGG GTAACAGTCC GTAAAGTTCC TGCCTCCTCG GACTTATCCG
V33cp    AATTTGATTC TACCGTGTGG GTGACAGTCC GTAAAGTTCC TGCCTCCTCG GACTTATCCG
Cmvv34   AATTTGATTC TACCGTGTGG GTGACAGTTC GTAAAGTTCC TGCCTCCTCG GACTTATCCG
Ccp      AATTTGATTC TACCGTGTGG GTGACAGTCC GTAAAGTTCC TGCCTCCTCG GACTTATCCG
Cmvwl    AATTTGATTC TACCGTGTGG GTTACAGTTC GGAAAGTACC TTCATCATCC GATCTTTCCG

         781                                                        840
V27cp    TTGCCGCCAT CTCTGCTATG TTCGCGGACG GAGCCTCACC GGTACTGGTT TATCAGTATG
V33cp    TTGCCGCCAT CTCTGCTATG TTTGCGGACG GAGCCTCACC GGTACTGGTT TATCAGTACG
Cmvv34   TTGCCGCCAT CTCTGCTATG TTCGCGGACG GAGCCTCACC GGTACTGGTT TATCAGTATG
Ccp      TTGCCGCCAT CTCTGCTATG TTCGCGGACG GAGCCTCACC GGTACTGGTT TATCAGTATG
Cmvwl    TCGCCGCCAT CTCTGCTATG TTTGGCGATG GTAATTCACC GGTTTTGGTT TATCAGTATG
```

FIG. 4C

```
           841                                                          900
 V27cp     CTGCATCTGG AGTCCAAGCT AACAACAAAT TGTTGTATGA TCTTTCGGCG ATGCGCGCTG
 V33cp     CTGCATCTGG AGTCCAAGCT AACAACAAAT TGTTGTATGA TCTTTCGGCG ATGCGCGCTG
Cmvv34     CTGCATCTGG AGTTCAAGCT AACAACAAAT TGTTGTATGA TCTTTCGGCG ATGCGCGCTG
   Ccp     CCGCATCTGG AGTCCAAGCC AACAACAAAC TGTTGTTTGA TCTTTCGGCG ATGCGCGCTG
Cmvwl      CTGCGTCCGG AGTTCAGGCC AACAATAAGT TACTTTATGA CCTGTCCGAG ATGCGTGCTG

           901                                                          960
 V27cp     ATATAGGTGA CATGAGAAAG TACGCCGTCC TCGTGTATTC AAAAGACGAT GCGCTCGAGA
 V33cp     ATATAGGCGA CATGAGAAAG TACGCCGTCC TCGTGTATTC AAAAGACGAT GCACTCGAGA
Cmvv34     ATATAGGTGA CATGAGAAAG TACGCCGTCC TCGTGTATTC AAAAGACGAT GCACTCGAGA
   Ccp     ATATAGGTGA CATGAGAAAG TACGCCGTCC TCGTGTATTC AAAAGACGAT GCGCTCGAGA
Cmvwl      ATATCGGCGA CATGCGTAAG TACGCCGTCC TGGTTTACTC GAAAGACGAT AAACTAGAGA

           961                                                         1020
 V27cp     CGGACGAGCT AGTACTTCAT GTTGACATCG AGCACCAACG TATTCCCACG TCTGGGATGC
 V33cp     CGGACGAGCT AGTACTTCAT GTTGACGTCG AGCACCAACG CATTCCCACG TCTGGGGTGC
Cmvv34     CGGACGAGCT AGTACTTCAT GTTGACATCG AGCACCAACG CATTCCCACG TCTGGGGTGC
   Ccp     CGGACGAGCT AGTACTTCAT GTTGACATCG AGCACCAACG CATTCCCACA TCTGGAGTGC
Cmvwl      AGGACGAGAT TGCACTTCAT GTCGACGTCG AGCATCAACG AATTCCTATC TCACGGATGC

           1021                                                        1080
 V27cp     TCC....... ..CAGTCTGA TTCCGTG.TT CCCAGAACCC T.CCCTCCGA TTTCTGTGGC
 V33cp     TCC....... ..CAGTATAA TTCTGTGCTT TCCAGAACCC T.CCCTCCGA TTTCTGTGGC
Cmvv34     TCC....... ..CAGTTTGA TTCCGTG.TT .CCAGAACCC T.CCCTCCGA TTTCTGTGGC
   Ccp     TCC....... ..CAGTCTGA TTCCGTG.TT CCCAGAACCC T.CCCTCCGA TCTCTGTGGC
Cmvwl      TCCCGACTTA GTCCGTGTGT TTACCGGCGT CCGAGAACGT TAAACTACAC TCTCAATCGC
```

FIG. 4D

```
             1081                                                        1140
  V27cp      GGGAGCTGAG TTGGCAGTTC TGCTATAAAC TGTCTGAAGT CACTAAACGT ....TTCACG
  V33cp      GGGAGCTGAG TTGGCAGTTC TGCTGTAAAC TGTCTGAAGT CACTAAACGT ....TTTACG
 Cmvv34      GGGAGCTGAG TTGGCAGTTC TGCTATAAAC TGTCTGAAGT CACTAAACGT ....TTTACG
    Ccp      GGGAGCTGAG TTGGCAGTTC TACTACAAAC TGTCTGGAGT CACTAAACGT ....TTTACG
 Cmvwl       GAGTGCTGAC TTGGTAGTAT TGCTTCAAAC TGCCTGAAGT CCCTAAACGT GTTGTTGCGC

             1141                                                        1200
  V27cp      GTGAACGGGT TGTCCATGG
  V33cp      GTGAACGGGT TGTCCATGG
 Cmvv34      GTGAACGGGT TGTCCATGG
    Ccp      GTGAACGGGT TGTCCATCCA GCTTACGGCT
 Cmvwl       GGGGAACGGG TGTCCATCCA GCTTACGGCT
             RMM352-->3' CAGGTACCT CGAATGCCGAGCTCACCAG 5'
                             Nco I
```

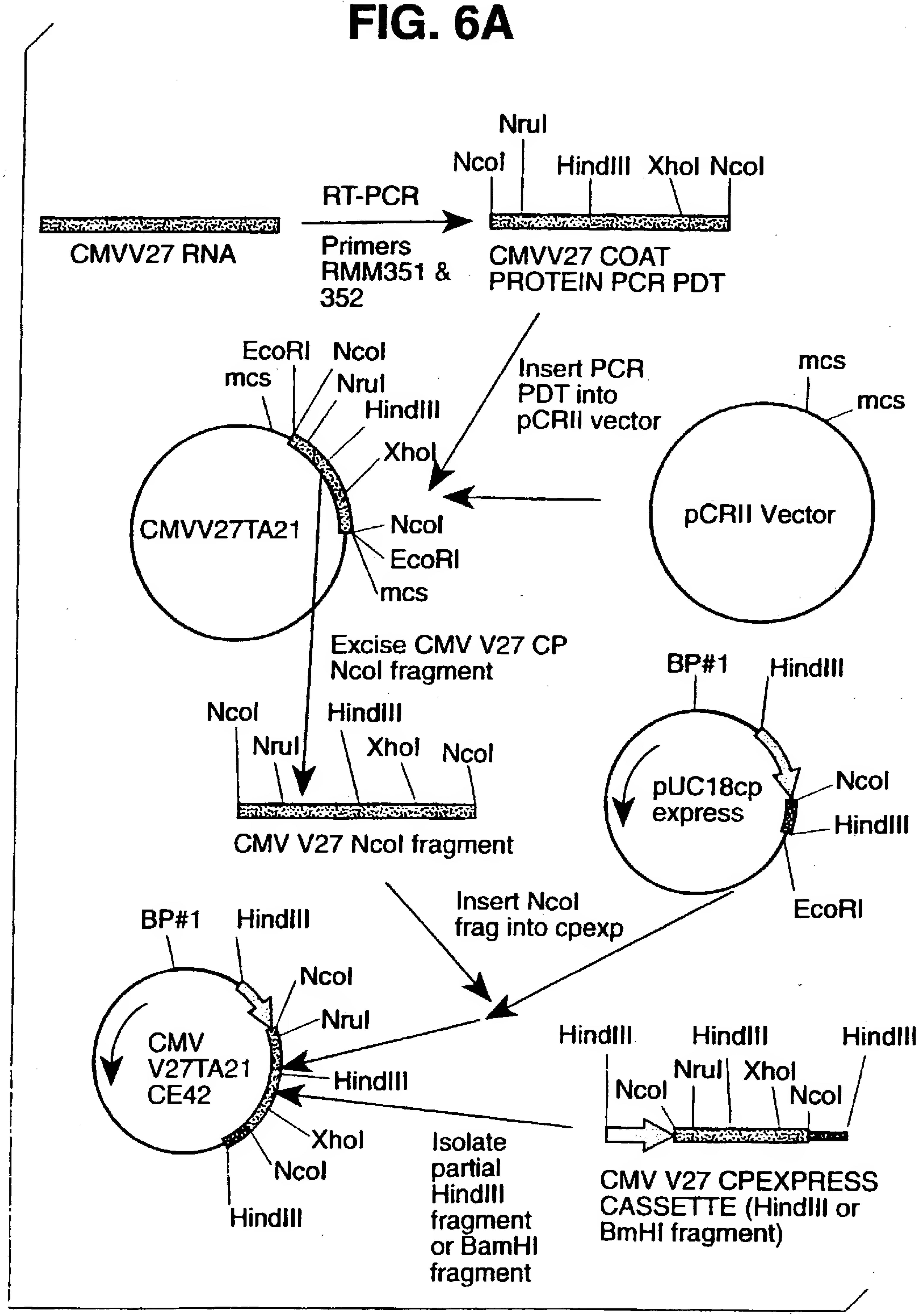
FIG. 5A

```
                1          *           *   *          *50
  Cmvv34  MDKSESTSAG R.NRRRRPRR GSRSASSSSD ANFRVLSQQL SRLNKTLAAG
  Cmvv27  MDKSESTSAG R.NRRRRPRR GSRSASSSSD ANFRVLSQQL SRLNKTLAAG
    Cmvc  MDKSESTSAG R.NHRRRPRR GSRSAPSSAD ANFRVLSQQL SRLNKTLAAG
   V33cp  MDKSESTSAG R.NRRRRPRR GSRSAPSSAD ANFRVLSQQL SRLNKTLSAG
   Cmvq3  MDKSGSPNAS RTSRRRRPRR GSRSA.SGAD AGLRALTQQM LRLNKTLAIG
   Cmvwl  MDKSGSPNAS RTSRRRRPRR GSRSA.SGAD AGLRALTQQM LKLNRTLAIG

          51              **                  *          100
  Cmvv34  RPTINHPTFV GSERCRPGYT FTSITLKPPK IDRGSYYGKR LLLPDSVTEY
  Cmvv27  RPTINHPTFV GSERCKPGYT FTSITLKPPK IDRGSYYGKR LLLPDSVTEY
    Cmvc  RPTINHPTFV GSERCRPGYT FTSITLKPPK IDRESYYGKR LLLPDSVTEY
   V33cp  RPTINHPTFV GSERCKSGYT FTSITLKPPK IDRGSYYGKR LLLPDSVTEY
   Cmvq3  RPTLNHPTFV GSESCKPGYT FTSITLKPPE IEKGSYFGRR LSLPDSVTDY
   Cmvwl  RPTLNHPTFV GSESCKPGYT FTSITLKPPE IEKGSYFGRR LSLPDSVTDY

          101                                            150
  Cmvv34  DKKLVSRIQI RVNPLPKFDS TVWVTVRKVP ASSDLSVAAI SAMFADGASP
  Cmvv27  DKKLVSRIQI RVNPLPKFDS TVWVTVRKVP ASSDLSVAAI SAMFADGASP
    Cmvc  DKKLVSRIQI RVNPLPKFDS TVWVTVRKVP ASSDLSVAAI SAMFADGASP
   V33cp  DKKLVSRIQI RVNPLPKFDS TVWVTVRKVP ASSDLSVAAI SAMFADGASP
   Cmvq3  DKKLVSRIQI RINPLPKFDS TVWVTVRKVP SSSDLSVAAI SAMFGDGNSP
   Cmvwl  DKKLVSRIQI RVNPLPKFDS TVWVTVRKVP SSSDLSVAAI SAMFGDGNSP
```

FIG. 5B

```
        151               *                                   200
Cmvv34  VLVYQYAASG VQANNKLLYD LSAMRADIGD MRKYAVLVYS KDDALETDEL
Cmvv27  VLVYQYAASG VQANNKLLYD LSAMRADIGD MRKYAVLVYS KDDALETDEL
  Cmvc  VLVYQYAASG VQANNKLLFD LSAMRADIGD MRKYAVLVYS KDDALETDEL
 V33cp  VLVYQYAASG VQANNKLLYD LSAMRADIGD MRKYAVLVYS KDDALETDEL
 Cmvq3  VLVYQYAASG VQANNKLLYD LSEMRADIGD MRKYAVLVYS KDDKLEKDEI
 Cmvwl  VLVYQYAASG VQANNKLLYD LSEMRADIGD MRKYAVLVYS KDDKLEKDEI

        201   *          *                                    250
Cmvv34  VLHVDIEHQR IPTSGVLPV*
Cmvv27  VLHVDIEHQR IPTSGMLPV*
  Cmvc  VLHVDIEHQR IPTSGVLPV*
 V33cp  VLHVDVEHQR IPTSGVLPV*
 Cmvq3  VLHVDVEHQR IPISRMLPT*
 Cmvwl  ALHVDVEHQR IPISRMLPT*
```

FIG. 6A



NruI
NcoI
HindIII
XhoI
NcoI
RT-PCR
CMVV27 RNA
Primers RMM351 & 352
CMVV27 COAT PROTEIN PCR PDT
EcoRI
NcoI
mcs
NruI
HindIII
XhoI
Insert PCR PDT into pCRII vector
mcs
mcs
CMVV27TA21
NcoI
EcoRI
mcs
pCRII Vector
Excise CMV V27 CP NcoI fragment
BP#1
HindIII
NcoI
HindIII
NruI
XhoI
NcoI
pUC18cp express
NcoI
HindIII
CMV V27 NcoI fragment
EcoRI
BP#1
HindIII
Insert NcoI frag into cpexp
NcoI
NruI
CMV V27TA21 CE42
HindIII
HindIII
HindIII
HindIII
NcoI
NruI
XhoI
NcoI
XhoI
NcoI
HindIII
Isolate partial HindIII fragment or BamHI fragment
CMV V27 CPEXPRESS CASSETTE (HindIII or BmHI fragment)

FIG. 6B

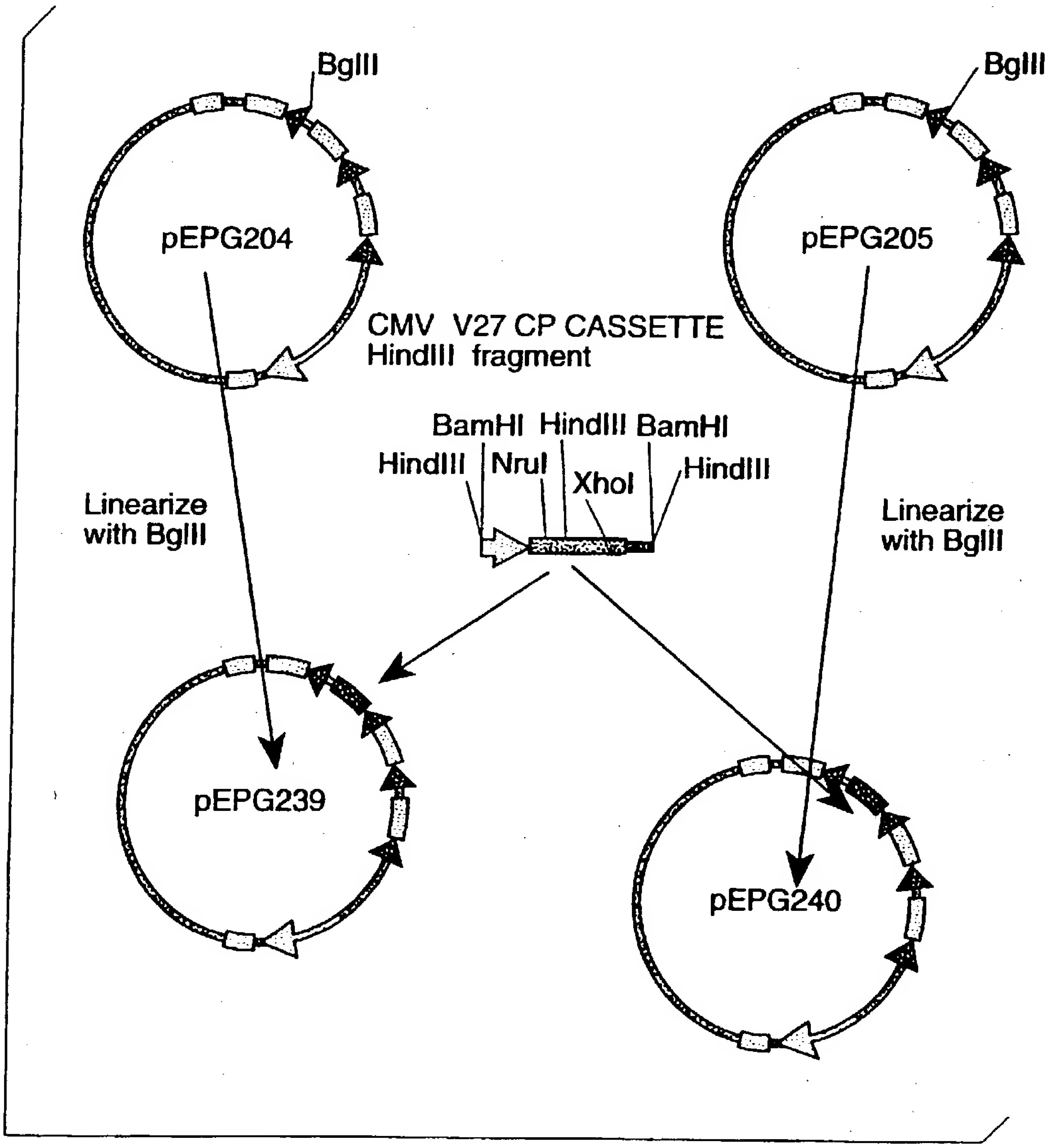
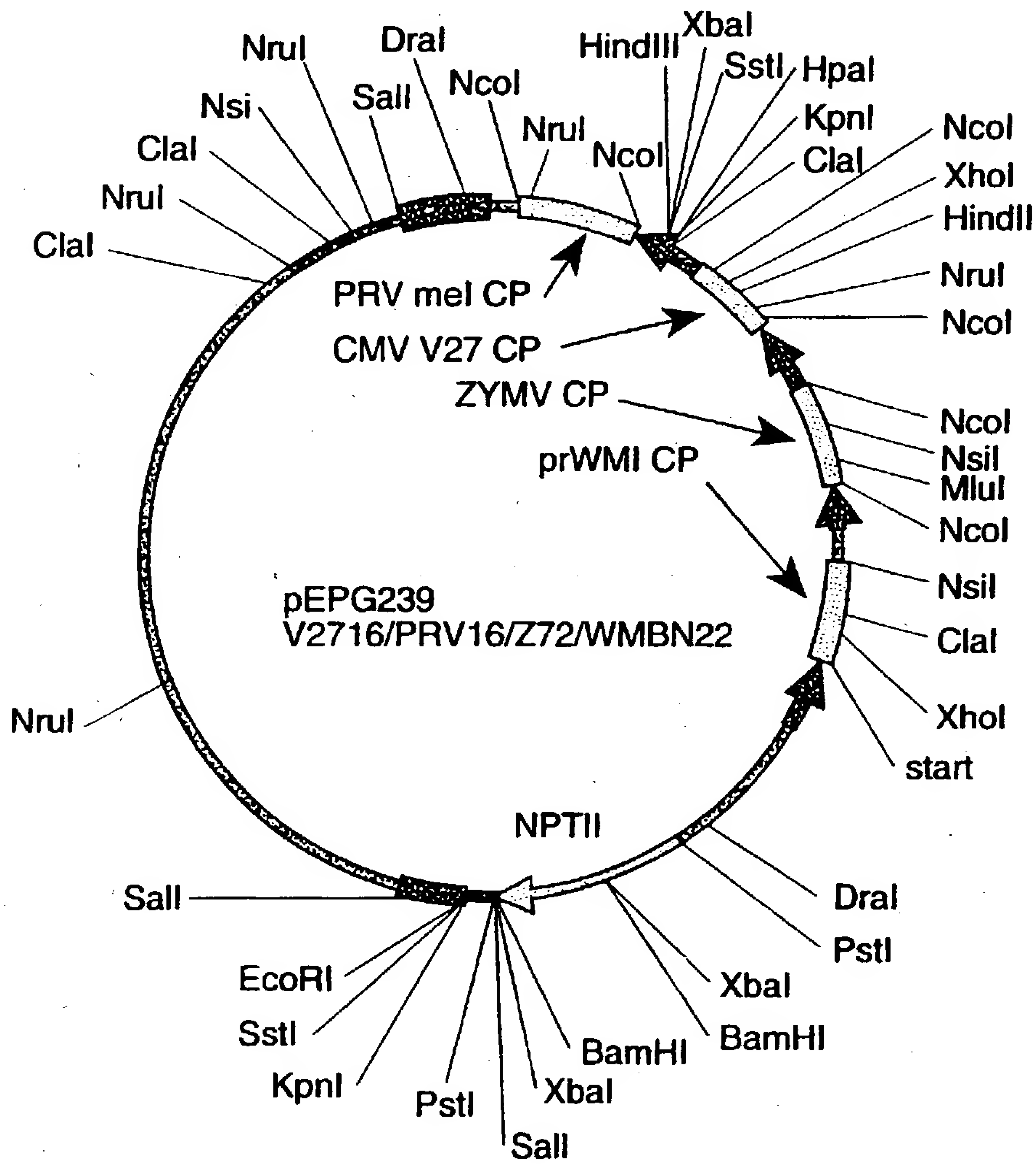

BglII
pEPG204
BglII
pEPG205
CMV V27 CP CASSETTE
HindIII fragment
BamHI HindIII BamHI
HindIII NruI
XhoI
HindIII
Linearize
with BglII
Linearize
with BglII
pEPG239
pEPG240

FIG. 6C



Nrul
Dral
HindIII
Xbal
SstI
Hpal
Nsi
Sall
Ncol
Kpnl
Nrul
Ncol
Ncol
Clal
Clal
Xhol
Nrul
HindIII
Clal
Nrul
PRV mel CP
Ncol
CMV V27 CP
ZYMV CP
Ncol
Nsil
prWMI CP
Mlul
Ncol
Nsil
pEPG239
V2716/PRV16/Z72/WMBN22
Clal
Nrul
Xhol
start
NPTII
Sall
Dral
Pstl
EcoRI
Xbal
SstI
BamHI
BamHI
Kpnl
Xbal
Pstl
Sall

FIG. 6D

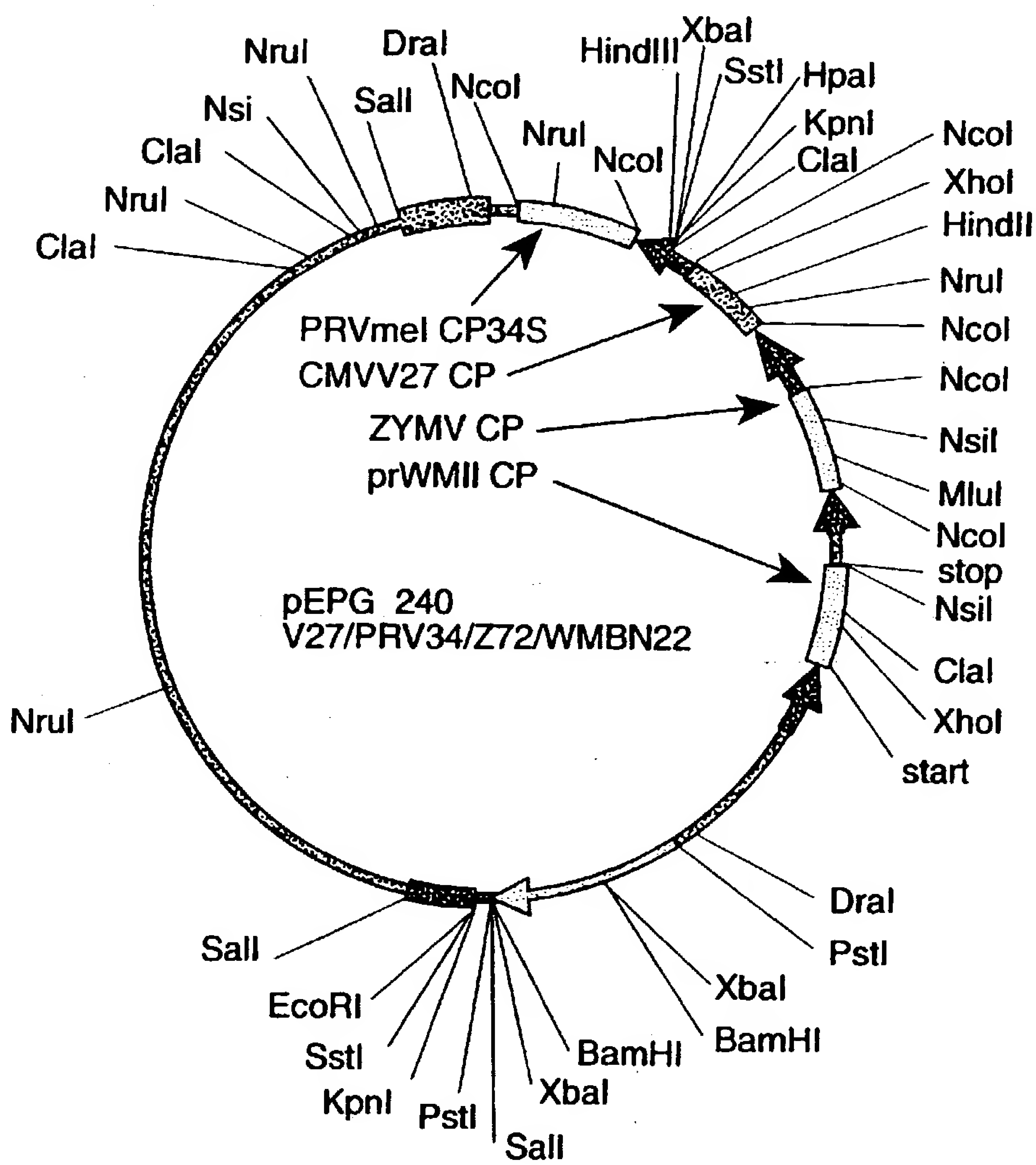
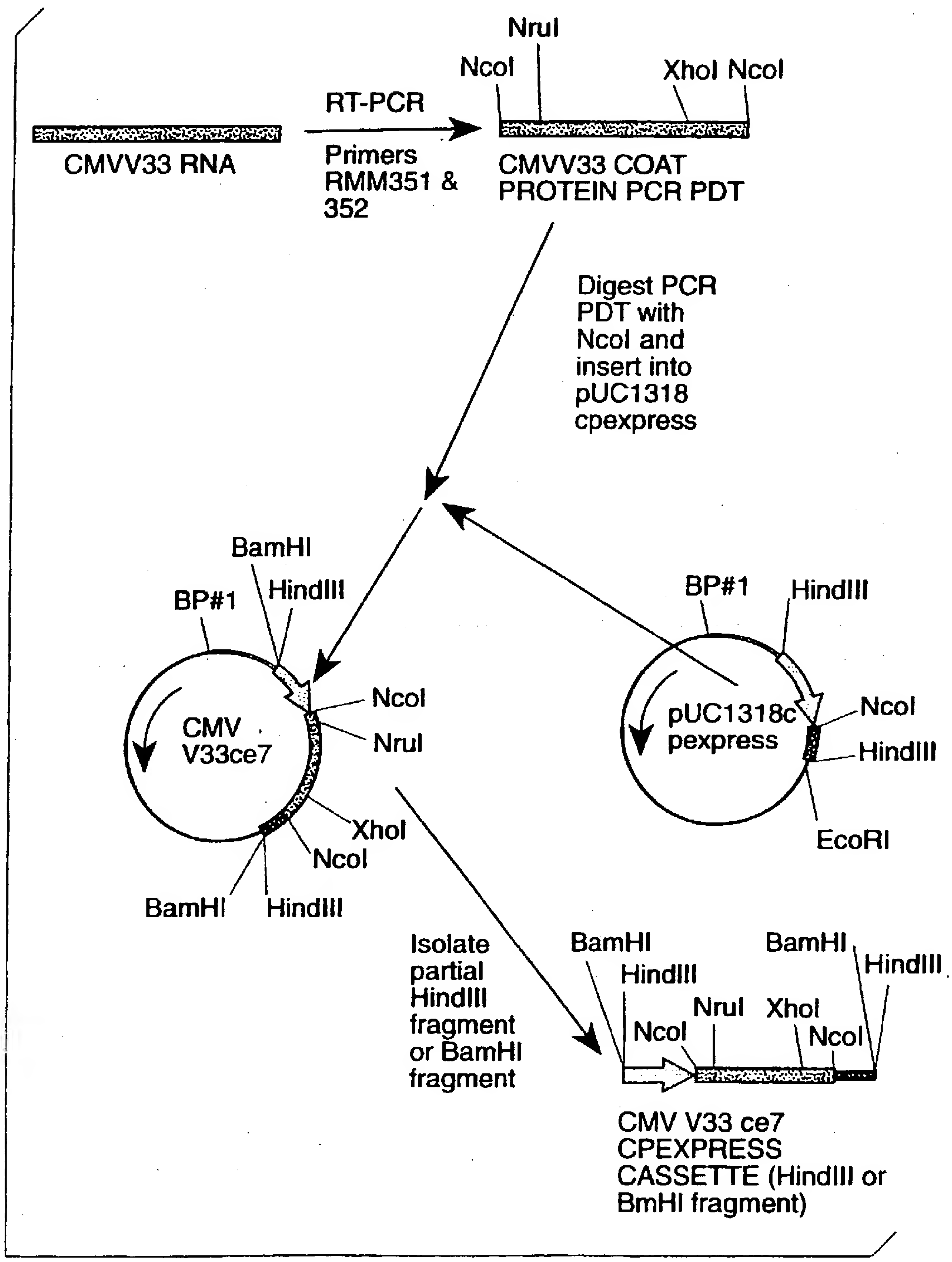

NruI
DraI
SalI
NcoI
Nsi
ClaI
NruI
ClaI
NruI
NcoI
HindIII
XbaI
SstI
HpaI
KpnI
ClaI
NcoI
XhoI
HindIII
NruI
NcoI
NcoI
NsiI
MluI
NcoI
stop
NsiI
ClaI
XhoI
start
PRVmel CP34S
CMVV27 CP
ZYMV CP
prWMII CP
pEPG 240
V27/PRV34/Z72/WMBN22
NruI
DraI
PstI
XbaI
BamHI
SalI
EcoRI
SstI
KpnI
PstI
SalI
XbaI
BamHI

FIG. 7A



NruI
NcoI
XhoI NcoI
RT-PCR
CMVV33 RNA
Primers
RMM351 &
352
CMVV33 COAT
PROTEIN PCR PDT
Digest PCR
PDT with
NcoI and
insert into
pUC1318
cpexpress
BamHI
BP#1
HindIII
CMV
V33ce7
NcoI
NruI
XhoI
NcoI
BamHI
HindIII
BP#1
HindIII
pUC1318c
pexpress
NcoI
HindIII
EcoRI
Isolate
partial
HindIII
fragment
or BamHI
fragment
BamHI
HindIII
NruI
XhoI
NcoI
BamHI
HindIII
NcoI
CMV V33 ce7
CPEXPRESS
CASSETTE (HindIII or
BmHI fragment)

FIG. 7B

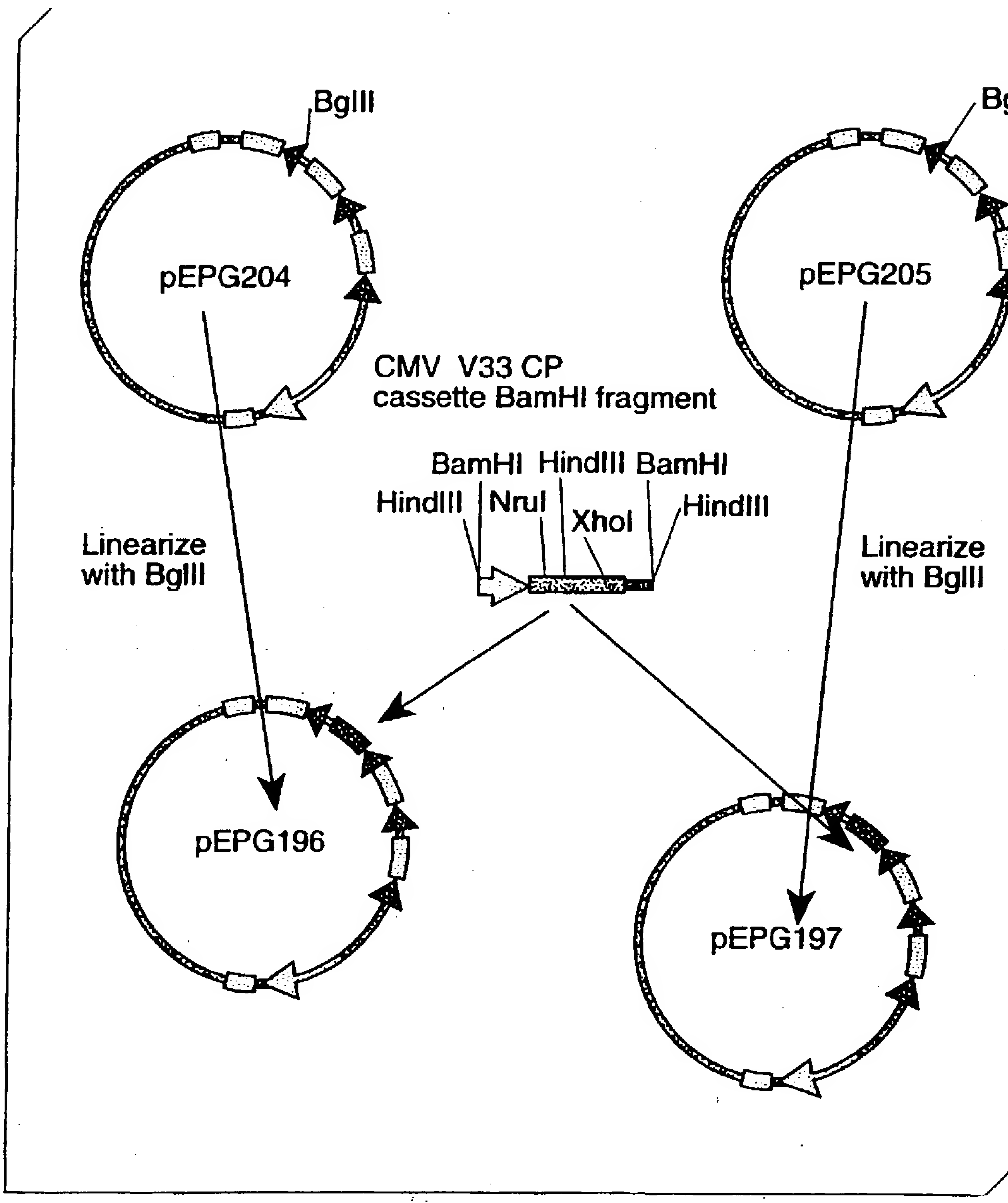
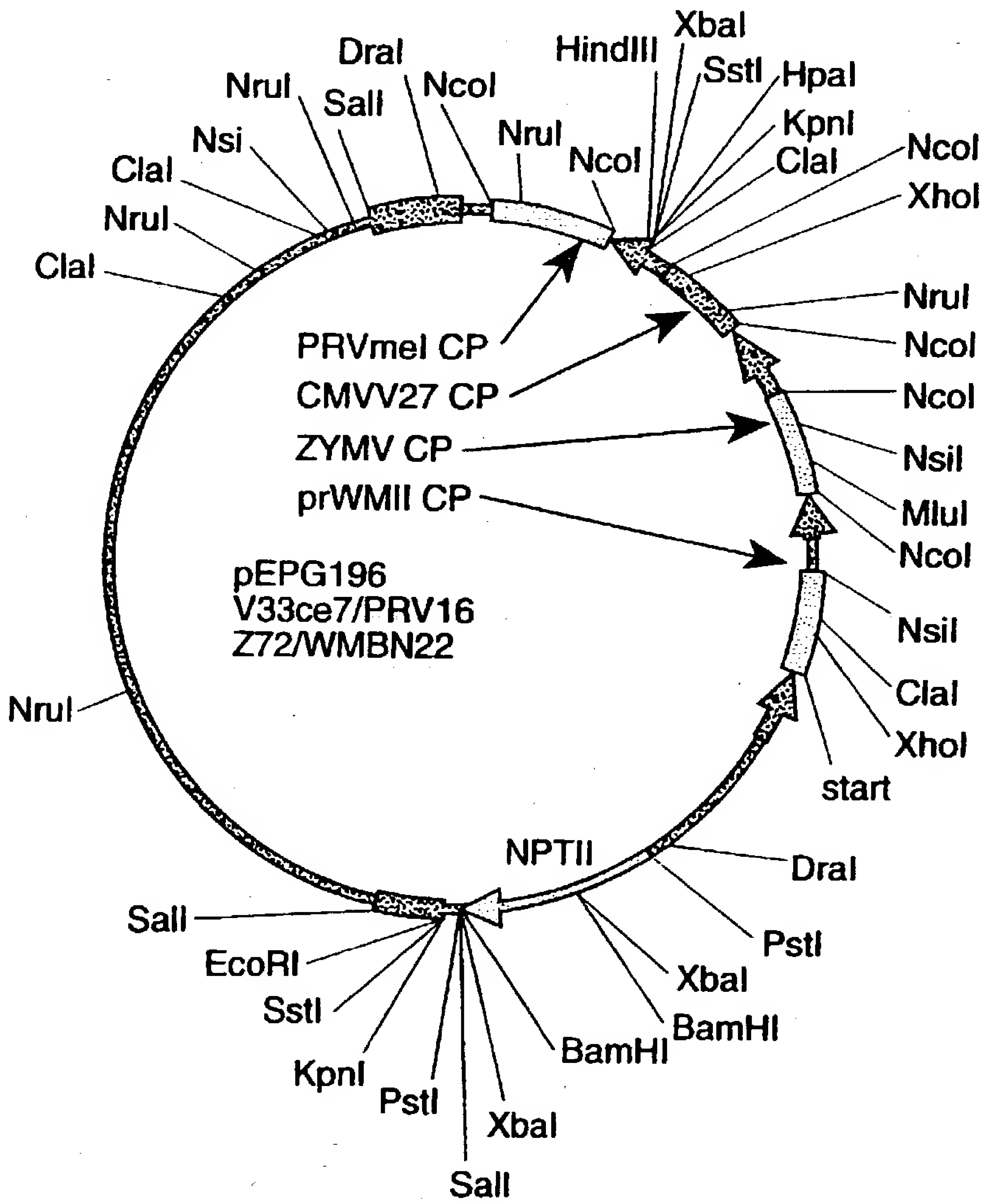
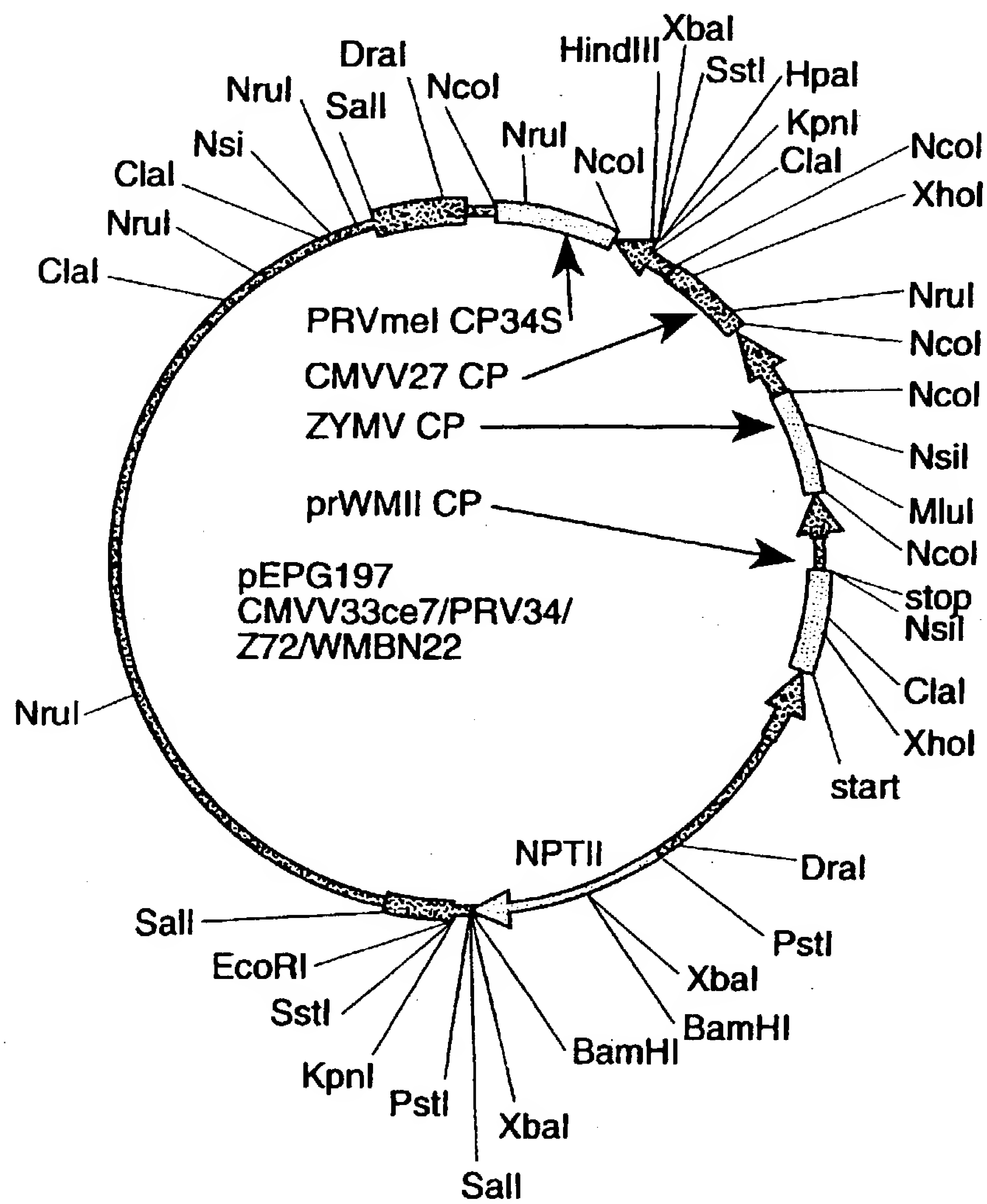

BglII
pEPG204
BglII
pEPG205
CMV V33 CP
cassette BamHI fragment
BamHI HindIII BamHI
HindIII NruI XhoI HindIII
Linearize
with BglII
Linearize
with BglII
pEPG196
pEPG197

FIG. 7C



Dral
Nrul
Sall
Ncol
HindIII
Xbal
Sstl
Hpal
Nsi
Nrul
Ncol
Kpnl
Clal
Clal
Ncol
Xhol
Nrul
Clal
Nrul
PRVmel CP
Ncol
CMVV27 CP
Ncol
ZYMV CP
Nsil
prWMII CP
Mlul
Ncol
pEPG196
V33ce7/PRV16
Z72/WMBN22
Nsil
Clal
Nrul
Xhol
start
NPTII
Dral
Sall
Pstl
EcoRI
Xbal
Sstl
BamHI
BamHI
Kpnl
Pstl
Xbal
Sall

FIG. 7D



XbaI
HindIII
SstI
HpaI
KpnI
ClaI
NcoI
XhoI
DraI
NruI
SalI
NcoI
NruI
NcoI
Nsi
ClaI
NruI
ClaI
PRVmel CP34S
CMVV27 CP
ZYMV CP
prWMII CP
pEPG197
CMVV33ce7/PRV34/
Z72/WMBN22
NruI
NcoI
NcoI
NsiI
MluI
NcoI
stop
NsiI
ClaI
XhoI
start
NruI
NPTII
DraI
PstI
XbaI
BamHI
BamHI
SalI
EcoRI
SstI
KpnI
PstI
XbaI
SalI

FIG. 8

```
CCATGGACAAATCTGAATCAACCAGTGCTGGTCGTAACCGTCGACGTCGTCCGCGTCGTG 60
A  M  D  K  S  E  S  T  S  A  G  R  N  R  R  R  R  P  R  R
GTTCCCGCTCCGCCCTCTCCTCCGCGGATGCTAACTTTAGAGTCCTGTCGCAGCAGCTTT 120
G  S  R  S  A  L  S  S  A  D  A  N  F  R  V  L  S  Q  Q  L
CGCGACTTAATAAGACGTTAGCAGCTGGTCGTCCAACTATTAACCACCCAACCTTTGTAG 180
S  R  L  N  K  T  L  A  A  G  R  P  T  I  N  H  P  T  F  V
GGAGTGAACGCTGTAGACCTGGGTACACGTTCACATCTATTACCCTAAAGCCACCAAAAA 240
G  S  E  R  C  R  P  G  Y  T  F  T  S  I  T  L  K  P  P  K
TAGACCGTGGGTCTTATTACGGTAAAAGGTTGTTACTACCTGATTCAGTCACAGAATATG 300
I  D  R  G  S  Y  Y  G  K  R  L  L  L  P  D  S  V  T  E  Y
ATAAGAAGCTTGTTTCGCGCATTCAAATTCGAGTTAATCCTTTGCCGAAATTTGATTCTA 360
D  K  K  L  V  S  R  I  Q  I  R  V  N  P  L  P  K  F  D  S
CCGTGTGGGTGACAGTCCGTAAAGTTCCTGCCTCCTCGGACTTATCCGTTGCCGCCATCT 420
T  V  W  V  T  V  R  K  V  P  A  S  S  D  L  S  V  A  A  I
CTGCTATGTTCGCGGACGGAGCCTCACCGGTACTGGTTTATCAGTATGCCGCATCTGGAG 480
S  A  M  F  A  D  G  A  S  P  V  L  V  Y  Q  Y  A  A  S  G
TCCAAGCCAACAACAAACTGTTGTATGATCTTTCGGCGATGCGCGCTGATATAGGTGACA 540
V  Q  A  N  N  K  L  L  Y  D  L  S  A  M  R  A  D  I  G  D
TGAGAAAGTACGCCGTCCTCGTGTATTCAAAAGACGATGCGCTCGAGACGGACGAGCTAG 600
M  R  K  Y  A  V  L  V  Y  S  K  D  D  A  L  E  T  D  E  L
TACTTCATGTTGACATCGAGCACCAACGCATTCCCACGTCTGGAGTGCTCCCAGTCTGAT 660
V  L  H  V  D  I  E  H  Q  R  I  P  T  S  G  V  L  P  V  .
TCTGTGTTCCCAGAACCCTCCCTCCGATCTCTGTGGCGGGAGCTGAGTTGGCAGTTCTGC 720
F  C  V  P  R  T  L  P  P  I  S  V  A  G  A  E  L  A  V  L
TGTAAACTGTCTGAAGTCACTAAACGTTTTACGGTGAACGGGTTGTCCATGG         772
L  .  T  V  .  S  H  .  T  F  Y  G  E  R  V  V  H  G
```

FIG. 9A

```
      M D K S E S T S A G R - N R R R R P R R G S R S A S S S A D A N F R V L S Q Q L  Majority
                        10                  20                  30                  40
  1   M D K S E S T S A G R - N H R R R P R R G S R S A P S S A D A N F R V L S Q Q L  CMV C AA SEQ
  1   M D K S E S T S A G R - N R R R R P R R G S R S A L S S A D A N F R V L S Q Q L  CMV CARNA5 AA SEQ
  1   M D K S E S T S A G R - N R R R R P R R G S R S A S S S S D A N F R V L S Q Q L  CMV V27 AA SEQ
  1   M D K S E S T S A G R - N R R R R P R R G S R S A P S S A D A N F R V L S Q Q L  CMV V33 AA SEQ
  1   M D K S E S T S A G R - N R R R R P R R G S R S A S S S S D A N F R V L S Q Q L  CMV V34 AA SEQ
  1   M D K S G S P N A S R T S R R R R P R R G S R S A - S G A D A G L R A L T Q Q M  CMV WL AA SEQ

      S R L N K T L A A G R P T I N H P T F V G S E R C K P G Y T F T S I T L K P P K  Majority
                        50                  60                  70                  80
 40   S R L N K T L A A G R P T I N H P T F V G S E R C R P G Y T F T S I T L K P P K  CMV C AA SEQ
 40   S R L N K T L A A G R P T I N H P T F V G S E R C R P G Y T F T S I T L K P P K  CMV CARNA5 AA SEQ
 40   S R L N K T L A A G R P T I N H P T F V G S E R C K P G Y T F T S I T L K P P K  CMV V27 AA SEQ
 40   S R L N K T L S A G R P T I N H P T F V G S E R C K S G Y T F T S I T L K P P K  CMV V33 AA SEQ
 40   S R L N K T L A A G R P T I N H P T F V G S E R C R P G Y T F T S I T L K P P K  CMV V34 AA SEQ
 40   L K L N R T L A I G R P T L N H P T F V G S E S C K P G Y T F T S I T L K P P E  CMV WL AA SEQ

      I D R G S Y Y G K R L L L P D S V T E Y D K K L V S R I Q I R V N P L P K F D S  Majority
                        90                 100                 110                 120
 80   I D R E S Y Y G K R L L L P D S V T E Y D K K L V S R I Q I R V N P L P K F D S  CMV C AA SEQ
 80   I D R G S Y Y G K R L L L P D S V T E Y D K K L V S R I Q I R V N P L P K F D S  CMV CARNA5 AA SEQ
 80   I D R G S Y Y G K R L L L P D S V T E Y D K K L V S R I Q I R V N P L P K F D S  CMV V27 AA SEQ
 80   I D R G S Y Y G K R L L L P D S V T E Y D K K L V S R I Q I R V N P L P K F D S  CMV V33 AA SEQ
 80   I D R G S Y Y G K R L L L P D S V T E Y D K K L V S R I Q I R V N P L P K F D S  CMV V34 AA SEQ
 80   I E K G S Y F G R R L S L P D S V T D Y D K K L V S R I Q I R V N P L P K F D S  CMV WL AA SEQ
```

FIG. 9B

```
     T V W V T V R K V P A S S D L S V A A I S A M F A D G A S P V L V Y Q Y A A S G  Majority
                       130                           140                           150                           160
120  T V W V T V R K V P A S S D L S V A A I S A M F A D G A S P V L V Y Q Y A A S G  CMV C AA SEQ
120  T V W V T V R K V P A S S D L S V A A I S A M F A D G A S P V L V Y Q Y A A S G  CMV CARNA5 AA SEQ
120  T V W V T V R K V P A S S D L S V A A I S A M F A D G A S P V L V Y Q Y A A S G  CMV V27 AA SEQ
120  T V W V T V R K V P A S S D L S V A A I S A M F A D G A S P V L V Y Q Y A A S G  CMV V33 AA SEQ
120  T V W V T V R K V P A S S D L S V A A I S A M F A D G A S P V L V Y Q Y A A S G  CMV V34 AA SEQ
120  T V W V T V R K V P[S]S S D L S V A A I S A M F[G]D G[N]S P V L V Y Q Y A A S G  CMV WL AA SEQ

     V Q A N N K L L Y D L S A M R A D I G D M R K Y A V L V Y S K D D A L E T D E L  Majority
                       170                           180                           190                           200
160  V Q A N N K L L[F]D L S A M R A D I G D M R K Y A V L V Y S K D D A L E T D E L  CMV C AA SEQ
160  V Q A N N K L L Y D L S A M R A D I G D M R K Y A V L V Y S K D D A L E T D E L  CMV CARNA5 AA SEQ
160  V Q A N N K L L Y D L S A M R A D I G D M R K Y A V L V Y S K D D A L E T D E L  CMV V27 AA SEQ
160  V Q A N N K L L Y D L S A M R A D I G D M R K Y A V L V Y S K D D A L E T D E L  CMV V33 AA SEQ
160  V Q A N N K L L Y D L S A M R A D I G D M R K Y A V L V Y S K D D A L E T D E L  CMV V34 AA SEQ
160  V Q A N N K L L Y D L S[E]M R A D I G D M R K Y A V L V Y S K D D[K]L E[K]D E[I]  CMV WL AA SEQ

     V L H V D I E H Q R I P T S G V L P V -                                          Majority
                       210                           220
200  V L H V D I E H Q R I P T S G V L P V .                                          CMV C AA SEQ
200  V L H V D I E H Q R I P T S G V L P V .                                          CMV CARNA5 AA SEQ
200  V L H V D I E H Q R I P T S G[M]L P V .                                          CMV V27 AA SEQ
200  V L H V D[V]E H Q R I P T S G V L P V .                                          CMV V33 AA SEQ
200  V L H V D I E H Q R I P T S G V L P V .                                          CMV V34 AA SEQ
200 [A]L H V D[V]E H Q R I P[I]S[R M]L P[T].                                          CMV WL AA SEQ
```

FIG. 10A

```
      X X X X X X X X X X X X X X X X X X X X   Majority
                        330                340
1     . . . . . . . . . . . . . . . . . . . .   carna5 cp cpexp33.seq
321   T A G A G A G T G T G T G T G C T G T G   New ccp.seq15
1     . . . . . . . . . . . . . . . . . . . .   New cmvv34.seq5
247   . . . . . . . . . T G A G T C G T G T G   New cmvwl.seq1
1     . . . . . . . . . . . . . . . . . . . .   New v27cp.seq5
1     . . . . . . . . . . . . . . . . . . . .   New v33cp.seq8

      X X X X X X X X X X X X X X X X X X X X   Majority
                        350                360
1     . . . . . . . . . . . . . . . . . . . .   carna5 cp cpexp33.seq
341   T T T T C T C T T T T G T G T C G T A G   New ccp.seq15
1     . . . . . . . . . . . . . . . . . . . .   New cmvv34.seq5
258   T T T T G T A T T T T G C G T C T T A G   New cmvwl.seq1
1     . . . . . . . . . . . . . . . . . . . .   New v27cp.seq5
1     . . . . . . . . . . . . . . . . . . . .   New v33cp.seq8

      X X X X X X X X X X X X C C A T G G A C   Majority
                        370                380
1     . . . . . . . . . . . . C C A T G G A C   carna5 cp cpexp33.seq
361   A A T T G A G T C G A G T C A T G G A C   New ccp.seq15
1     . . . . . . . . . . . . C C A T G G A C   New cmvv34.seq5
278   . . . T G T G C . . . . C T A T G G A C   New cmvwl.seq1
1     . . . . . . . . . . . . C C A T G G A C   New v27cp.seq5
1     . . . . . . . . . . . . C C A T G G A C   New v33cp.seq8

      A A A T C T G A A T C A A C C A G T G C   Majority
                        390                400
9     A A A T C T G A A T C A A C C A G T G C   carna5 cp cpexp33.seq
381   A A A T C T G A A T C A A C C A G T G C   New ccp.seq15
9     A A A T C T G A A T C A A C C A G T G C   New cmvv34.seq5
291   A A A T C T G G A T C T C C C A A T G C   New cmvwl.seq1
9     A A A T C T G A A T C A A C C A G T G C   New v27cp.seq5
9     A A A T C T G A A T C A A C C A G T G C   New v33cp.seq8
```

FIG. 10B

```
    T G G T C G T A A C C G T C G A C G T C   Majority
                      410               420
29  T G G T C G T A A C C G T C G A C G T C   carna5 cp cpexp33.seq
401 T G G T C G T A A C C A T C G A C G T C   New ccp.seq15
29  T G G T C G T A A C C G T C G A C G T C   New cmvv34.seq5
311 T A G T A G A A C C T C C C G G C G T C   New cmvwl.seq1
29  T G G T C G T A A C C G T C G G C G T C   New v27cp.seq5
29  T G G T C G T A A C C G T C G A C G T C   New v33cp.seq8

    G T C X X X C G C G T C G T G G T T C C   Majority
                      430               440
49  G T C . . . C G C G T C G T G G T T C C   carna5 cp cpexp33.seq
421 G T C . . . C G C G T C G T G G T T C C   New ccp.seq15
49  G T C . . . C G C G T C G T G G T T C C   New cmvv34.seq5
331 G T C G C C C G C G T A G A G G T T C T   New cmvwl.seq1
49  G T C . . . C G C G T C G T G G T T C C   New v27cp.seq5
49  G T C . . . C G C G T C G T G G T T C C   New v33cp.seq8

    C G C T C C G C C C C C T C C T C C G C   Majority
                      450               460
66  C G C T C C G C C C T C T C C T C C G C   carna5 cp cpexp33.seq
438 C G C T C C G C C C C C T C C T C C G C   New ccp.seq15
66  C G C T C C G C T T C C T C C T C T T C   New cmvv34.seq5
351 C G G T C C G C T . . . T C T G G T G C   New cmvwl.seq1
66  C G C T C C G C C T C T T C C T C C T C   New v27cp.seq5
66  C G C T C C G C C C C C T C C T C C G C   New v33cp.seq8

    G G A T G C T A A C T T T A G A G T C T   Majority
                      470               480
86  G G A T G C T A A C T T T A G A G T C C   carna5 cp cpexp33.seq
458 G G A T G C T A A C T T T A G A G T C T   New ccp.seq15
86  G G A T G C T A A C T T T A G A G T C T   New cmvv34.seq5
368 G G A T G C A G G G T T G C G T G C T T   New cmvwl.seq1
86  G G A T G C T A A C T T T A G A G T C T   New v27cp.seq5
86  G G A T G C C A A C T T T A G A G T C T   New v33cp.seq8
```

FIG. 10C

```
    T G T C G C A G C A G C T T T C G C G A  Majority
                      490                500
106 T G T C G C A G C A G C T T T C G C G A  carna5 cp cpexp33.seq
478 T G T C G C A G C A G C T T T C G C G A  New ccp.seq15
106 T G T C G C A G C A G C T T T C G C G A  New cmvv34.seq5
388 T G A C T C A G C A G A T G C T G A A A  New cmvwl.seq1
106 T G T C G C A G C A G C T T T C G C G A  New v27cp.seq5
106 T G T C G C A G C A G C T T T C G C G A  New v33cp.seq8

    C T T A A T A A G A C G T T A G C A G C  Majority
                      510                520
126 C T T A A T A A G A C G T T A G C A G C  carna5 cp cpexp33.seq
498 C T T A A T A A G A C G T T A G C A G C  New ccp.seq15
126 C T T A A C A A G A C G T T A G C A G C  New cmvv34.seq5
408 C T C A A T A G A A C C C T C G C C A T  New cmvwl.seq1
126 C T T A A C A A G A C G T T A G C A G C  New v27cp.seq5
126 C T T A A T A A G A C G T T G T C A G C  New v33cp.seq8

    T G G T C G T C C A A C T A T T A A C C  Majority
                      530                540
146 T G G T C G T C C A A C T A T T A A C C  carna5 cp cpexp33.seq
518 T G G T C G T C C A A C T A T T A A C C  New ccp.seq15
146 T G G T C G T C C A A C T A T T A A C C  New cmvv34.seq5
428 T G G T C G T C C C A C T C T T A A C C  New cmvwl.seq1
146 T G G T C G T C C A A C T A T T A A C C  New v27cp.seq5
146 T G G T C G T C C A A C T A T T A A C C  New v33cp.seq8

    A C C C A A C C T T T G T A G G G A G T  Majority
                      550                560
166 A C C C A A C C T T T G T A G G G A G T  carna5 cp cpexp33.seq
538 A C C C A A C C T T T G T A G G G A G T  New ccp.seq15
166 A C C C A A C C T T T G T A G G G A G T  New cmvv34.seq5
448 A C C C A A C C T T C G T G G G T A G T  New cmvwl.seq1
166 A C C C A A C C T T T G T A G G G A G T  New v27cp.seq5
166 A C C C A A C C T T T G T A G G G A G T  New v33cp.seq8
```

FIG. 10D

```
      G A A C G C T G T A G A C C T G G G T A  Majority
                        570                 580
186   G A A C G C T G T A G A C C T G G G T A  carna5 cp cpexp33.seq
558   G A A C G C T G T A G A C C T G G G T A  New ccp.seq15
186   G A A C G C T G T A G A C C T G G G T A  New cmvv34.seq5
468   G A A A G C T G T A A A C C C G G T T A  New cmvwl.seq1
186   G A A C G C T G T A A A C C T G G G T A  New v27cp.seq5
186   G A G C G T T G T A A A T C T G G G T A  New v33cp.seq8

      C A C G T T C A C A T C T A T T A C C C  Majority
                        590                 600
206   C A C G T T C A C A T C T A T T A C C C  carna5 cp cpexp33.seq
578   C A C G T T C A C A T C T A T T A C C C  New ccp.seq15
206   C A C G T T C A C A T C T A T T A C C C  New cmvv34.seq5
488   C A C T T T C A C A T C T A T T A C C C  New cmvwl.seq1
206   C A C G T T C A C A T C T A T T A C C C  New v27cp.seq5
206   C A C G T T C A C A T C T A T T A C C C  New v33cp.seq8

      T A A A G C C A C C A A A A A T A G A C  Majority
                        610                 620
226   T A A A G C C A C C A A A A A T A G A C  carna5 cp cpexp33.seq
598   T A A A G C C A C C A A A A A T A G A C  New ccp.seq15
226   T A A A G C C A C C A A A A A T A G A C  New cmvv34.seq5
508   T G A A A C C G C C T G A A A T T G A G  New cmvwl.seq1
226   T A A A G C C A C C A A A A A T A G A C  New v27cp.seq5
226   T A A A G C C G C C G A A A A T A G A C  New v33cp.seq8

      C G T G G G T C T T A T T A C G G T A A  Majority
                        630                 640
246   C G T G G G T C T T A T T A C G G T A A  carna5 cp cpexp33.seq
618   C G T G A G T C T T A T T A C G G T A A  New ccp.seq15
246   C G C G G G T C T T A C T A C G G T A A  New cmvv34.seq5
528   A A A G G T T C A T A T T T T G G T A G  New cmvwl.seq1
246   C G T G G G T C T T A T T A C G G T A A  New v27cp.seq5
246   C G T G G G T C T T A T T A T G G T A A  New v33cp.seq8
```

FIG. 10E

	A A G G T T G T T A T T A C C T G A T T	Majority
	650 660	
266	A A G G T T G T T A C T A C C T G A T T	carna5 cp cpexp33.seq
638	A A G G T T G T T A C T A C C T G A T T	New ccp.seq15
266	A A G G T T G T T A C T A C C T G A T T	New cmvv34.seq5
548	A A G G T T G T C T T T G C C A G A T T	New cmvwl.seq1
266	A A G G T T G T T A T T A C C T G A T T	New v27cp.seq5
266	A A G G T T G T T A T T A C C T G A T T	New v33cp.seq8

	C A G T C A C G G A A T A T G A T A A G	Majority
	670 680	
286	C A G T C A C A G A A T A T G A T A A G	carna5 cp cpexp33.seq
568	C A G T C A C G G A A T A T G A T A A G	New ccp.seq15
286	C A G T C A C G G A A T A T G A T A A G	New cmvv34.seq5
568	C A G T C A C G G A C T A T G A T A A G	New cmvwl.seq1
286	C A G T C A C G G A A T A T G A T A A G	New v27cp.seq5
286	C A G T C A C A G A A T A T G A T A A G	New v33cp.seq8

	A A G C T T G T T T C G C G C A T T C A	Majority
	690 700	
306	A A G C T T G T T T C G C G C A T T C A	carna5 cp cpexp33.seq
678	A A G C T T G T T T C G C G C A T T C A	New ccp.seq15
306	A A G C T T G T T T C G C G C A T T C A	New cmvv34.seq5
588	A A G C T T G T T T C G C G C A T T C A	New cmvwl.seq1
306	A A G C T T G T T T C G C G C A T T C A	New v27cp.seq5
306	A A A C T T G T T T C G C G C A T T C A	New v33cp.seq8

	A A T T C G A G T T A A T C C T T T G C	Majority
	710 720	
326	A A T T C G A G T T A A T C C T T T G C	carna5 cp cpexp33.seq
698	A A T T C G A G T T A A T C C T T T G C	New ccp.seq15
326	A A T T C G A G T T A A T C C T T T G C	New cmvv34.seq5
608	A A T C A G G G T T A A T C C T T T G C	New cmvwl.seq1
326	A A T T C G A G T T A A T C C T T T G C	New v27cp.seq5
326	A A T T C G A G T T A A T C C C T T G C	New v33cp.seq8

FIG. 10F

	C G A A A T T T G A T T C T A C C G T G	Majority
	730 740	
346	C G A A A T T T G A T T C T A C C G T G	carna5 cp cpexp33.seq
718	C G A A A T T T G A T T C T A C C G T G	New ccp.seq15
346	C G A A A T T T G A T T C T A C C G T G	New cmvv34.seq5
628	C G A A A T T T G A T T C T A C C G T G	New cmvwl.seq1
346	C G A A A T T T G A T T C T A C C G T G	New v27cp.seq5
346	C G A A A T T T G A T T C T A C C G T G	New v33cp.seq8

	T G G G T G A C A G T C C G T A A A G T	Majority
	750 760	
266	T G G G T G A C A G T C C G T A A A G T	carna5 cp cpexp33.seq
738	T G G G T G A C A G T C C G T A A A G T	New ccp.seq15
366	T G G G T G A C A G T T C G T A A A G T	New cmvv34.seq5
648	T G G G T T A C A G T T C G G A A A G T	New cmvwl.seq1
366	T G G G T A A C A G T C C G T A A A G T	New v27cp.seq5
366	T G G G T G A C A G T C C G T A A A G T	New v33cp.seq8

	T C C T G C C T C C T C G G A C T T A T	Majority
	770 780	
386	T C C T G C C T C C T C G G A C T T A T	carna5 cp cpexp33.seq
758	T C C T G C C T C C T C G G A C T T A T	New ccp.seq15
386	T C C T G C C T C C T C G G A C T T A T	New cmvv34.seq5
668	A C C T T C A T C A T C C G A T C T T T	New cmvwl.seq1
386	T C C T G C C T C C T C G G A C T T A T	New v27cp.seq5
386	T C C T G C C T C C T C G G A C T T A T	New v33cp.seq8

	C C G T T G C C G C C A T C T C T G C T	Majority
	790 800	
406	C C G T T G C C G C C A T C T C T G C T	carna5 cp cpexp33.seq
778	C C G T T G C C G C C A T C T C T G C T	New ccp.seq15
406	C C G T T G C C G C C A T C T C T G C T	New cmvv34.seq5
688	C C G T C G C C G C C A T C T C T G C T	New cmvwl.seq1
406	C C G T T G C C G C C A T C T C T G C T	New v27cp.seq5
406	C C G T T G C C G C C A T C T C T G C T	New v33cp.seq8

FIG. 10G

```
     A T G T T C G C G G A C G G A G C C T C   Majority
                        810                820
426  A T G T T C G C G G A C G G A G C C T C   carna5 cp cpexp33.seq
798  A T G T T C G C G G A C G G A G C C T C   New ccp.seq15
426  A T G T T C G C G G A C G G A G C C T C   New cmvv34.seq5
708  A T G T T T G G C G A T G G T A A T T C   New cmvwl.seq1
426  A T G T T C G C G G A C G G A G C C T C   New v27cp.seq5
426  A T G T T T G C G G A C G G A G C C T C   New v33cp.seq8

     A C C G G T A C T G G T T T A T C A G T   Majority
                        830                840
446  A C C G G T A C T G G T T T A T C A G T   carna5 cp cpexp33.seq
818  A C C G G T A C T G G T T T A T C A G T   New ccp.seq15
446  A C C G G T A C T G G T T T A T C A G T   New cmvv34.seq5
728  A C C G G T T T T G G T T T A T C A G T   New cmvwl.seq1
446  A C C G G T A C T G G T T T A T C A G T   New v27cp.seq5
446  A C C G G T A C T G G T T T A T C A G T   New v33cp.seq8

     A T G C T G C A T C T G G A G T C C A A   Majority
                        850                860
466  A T G C C G C A T C T G G A G T C C A A   carna5 cp cpexp33.seq
838  A T G C C G C A T C T G G A G T C C A A   New ccp.seq15
466  A T G C T G C A T C T G G A G T T C A A   New cmvv34.seq5
748  A T G C T G C G T C C G G A G T T C A G   New cmvwl.seq1
466  A T G C T G C A T C T G G A G T C C A A   New v27cp.seq5
466  A C G C T G C A T C T G G A G T C C A A   New v33cp.seq8

     G C T A A C A A C A A A T T G T T G T A   Majority
                        870                880
486  G C C A A C A A C A A A C T G T T G T A   carna5 cp cpexp33.seq
858  G C C A A C A A C A A A C T G T T G T T   New ccp.seq15
486  G C T A A C A A C A A A T T G T T G T A   New cmvv34.seq5
768  G C C A A C A A T A A G T T A C T T T A   New cmvwl.seq1
486  G C T A A C A A C A A A T T G T T G T A   New v27cp.seq5
486  G C T A A C A A C A A A T T G T T G T A   New v33cp.seq8
```

FIG. 10H

```
      T G A T C T T T C G G C G A T G C G C G   Majority
                        890                900
506   T G A T C T T T C G G C G A T G C G C G   carna5 cp cpexp33.seq
878   T G A T C T T T C G G C G A T G C G C G   New ccp.seq15
506   T G A T C T T T C G G C G A T G C G C G   New cmvv34.seq5
788   T G A C C T G T C C G A G A T G C G T G   New cmvwl.seq1
506   T G A T C T T T C G G C G A T G C G C G   New v27cp.seq5
506   T G A T C T T T C G G C G A T G C G C G   New v33cp.seq8

      C T G A T A T A G G T G A C A T G A G A   Majority
                        910                920
526   C T G A T A T A G G T G A C A T G A G A   carna5 cp cpexp33.seq
898   C T G A T A T A G G T G A C A T G A G A   New ccp.seq15
526   C T G A T A T A G G T G A C A T G A G A   New cmvv34.seq5
808   C T G A T A T C G G C G A C A T G C G T   New cmvwl.seq1
526   C T G A T A T A G G T G A C A T G A G A   New v27cp.seq5
526   C T G A T A T A G G C G A C A T G A G A   New v33cp.seq8

      A A G T A C G C C G T C C T C G T G T A   Majority
                        930                940
546   A A G T A C G C C G T C C T C G T G T A   carna5 cp cpexp33.seq
918   A A G T A C G C C G T C C T C G T G T A   New ccp.seq15
546   A A G T A C G C C G T C C T C G T G T A   New cmvv34.seq5
828   A A G T A C G C C G T C C T G G T T T A   New cmvwl.seq1
546   A A G T A C G C C G T C C T C G T G T A   New v27cp.seq5
546   A A G T A C G C C G T C C T C G T G T A   New v33cp.seq8

      T T C A A A A G A C G A T G C G C T C G   Majority
                        950                960
566   T T C A A A A G A C G A T G C G C T C G   carna5 cp cpexp33.seq
938   T T C A A A A G A C G A T G C G C T C G   New ccp.seq15
566   T T C A A A A G A C G A T G C A C T C G   New cmvv34.seq5
848   C T C G A A A G A C G A - A A A C T A G   New cmvwl.seq1
566   T T C A A A A G A C G A T G C G C T C G   New v27cp.seq5
566   T T C A A A A G A C G A T G C A C T C G   New v33cp.seq8
```

FIG. 10I

```
     A G A C G G A C G A G C T A G T A C T T   Majority
                       970                980
586  A G A C G G A C G A G C T A G T A C T T   carna5 cp cpexp33.seq
958  A G A C G G A C G A G C T A G T A C T T   New ccp.seq15
586  A G A C G G A C G A G C T A G T A C T T   New cmvv34.seq5
868  A G A A G G A C G A G A T T G C A C T T   New cmvwl.seq1
586  A G A C G G A C G A G C T A G T A C T T   New v27cp.seq5
586  A G A C G G A C G A G C T A G T A C T T   New v33cp.seq8

     C A T G T T G A C A T C G A G C A C C A   Majority
                       990               1000
606  C A T G T T G A C A T C G A G C A C C A   carna5 cp cpexp33.seq
978  C A T G T T G A C A T C G A G C A C C A   New ccp.seq15
906  C A T G T T G A C A T C G A G C A C C A   New cmvv34.seq5
888  C A T G T C G A C G T C G A G C A T C A   New cmvwl.seq1
606  C A T G T T G A C A T C G A G C A C C A   New v27cp.seq5
606  C A T G T T G A C G T C G A G C A C C A   New v33cp.seq8

     A C G C A T T C C C A C G T C T G G G G   Majority
                      1010               1020
626  A C G C A T T C C C A C G T C T G G A G   carna5 cp cpexp33.seq
998  A C G C A T T C C C A C A T C T G G A G   New ccp.seq15
626  A C G C A T T C C C A C G T C T G G G G   New cmvv34.seq5
908  A C G A A T T C C T A T C T C A C G G A   New cmvwl.seq1
626  A C G T A T T C C C A C G T C T G G G A   New v27cp.seq5
626  A C G C A T T C C C A C G T C T G G G G   New v33cp.seq8

     T G C T C C C A G T C T G A T T C X T G   Majority
                      1030               1040
646  T G C T C C C A G T C T G A T T C . T G   carna5 cp cpexp33.seq
1018 T G C T C C C A G T C T   A T T C . C G   New ccp.seq15
646  T G C T C C C A G T T T G A T T C . C G   New cmvv34.seq5
928  T G C T C C C G A C T T A G T C C G T G   New cmvwl.seq1
646  T G C T C C C A G T C T G A T T C . C G   New v27cp.seq5
646  T G C T C C C A G T A T A A T T C . T G   New v33cp.seq8
```

FIG. 10J

```
       T G X T T C C C X X X X X X X X A G A A   Majority
                         1050                1060
665    T G . T T C C C . . . . . . . . A G A A   carna5 cp cpexp33.seq
1037   T G . T T C C C . . . . . . . . A G A A   New ccp.seq15
665    T G . T T C C . . . . . . . . . A G A A   New cmvv34.seq5
948    T G T T T A C C G G C G T C C G A G A A   New cmvwl.seq1
665    T G . T T C C C . . . . . . . . A G A A   New v27cp.seq5
665    T G C T T T C C . . . . . . . . A G A A   New v33cp.seq8
```

```
       C C C T C C X C T C C G A T T T C T G T   Majority
                         1070                1080
676    C C C T C C . C T C C G A T C T C T G T   carna5 cp cpexp33.seq
1048   C C C T C C . C T C C G A T C T C T G T   New ccp.seq15
675    C C C T C C . C T C C G A T T T C T G T   New cmvv34.seq5
968    C G T T A A A C T A C A C T C T C A A T   New cmvwl.seq1
676    C C C T C C . C T C C G A T T T C T G T   New v27cp.seq5
677    C C C T C C . C T C C G A T T T C T G T   New v33cp.seq8
```

```
       G G C G G G A G C T G A G T T G G C A G   Majority
                         1090                1100
695    G G C G G G A G C T G A G T T G G C A G   carna5 cp cpexp33.seq
1067   G G C G G G A G C T G A G T T G G C A G   New ccp.seq15
694    G G C G G G A G C T G A G T T G G C A G   New cmvv34.seq5
988    C G C G A G T G C T G A C T T G G T A G   New cmvwl.seq1
695    G G C G G G A G C T G A G T T G G C A G   New v27cp.seq5
696    G G C G G G A G C T G A G T T G G C A G   New v33cp.seq8
```

```
       T T C T G C T A T A A A C T G T C T G A   Majority
                         1110                1120
715    T T C T G C T G T A A A C T G T C T G A   carna5 cp cpexp33.seq
1087   T T C T A C T A C A A A C T G T C T G G   New ccp.seq15
714    T T C T G C T A T A A A C T G T C T G A   New cmvv34.seq5
1008   T A T T G C T T C A A A C T G C C T G A   New cmvwl.seq1
715    T T C T G C T A T A A A C T G T C T G A   New v27cp.seq5
716    T T C T G C T G T A A A C T G T C T G A   New v33cp.seq8
```

FIG. 10K

```
     A G T C A C T A A A C G T T T T A X X X  Majority
                       1130            1140
735  A G T C A C T A A A C G T T T T A . . .  carna5 cp cpexp33.seq
1107 A G T C A C T A A A C G T T T T A . . .  New ccp.seq15
734  A G T C A C T A A A C G T T T T A . . .  New cmvv34.seq5
1028 A G T C C C T A A A C G T G T T G T T G  New cmvwl.seq1
735  A G T C A C T A A A C G T T T C A . . .  New v27cp.seq5
736  A G T C A C T A A A C G T T T T A . . .  New v33cp.seq8

     X X C G G T G A A C G G G T T G T C C A  Majority
                       1150            1160
752  . . C G G T G A A C G G G T T G T C C A  carna5 cp cpexp33.seq
1124 . . C G G T G A A C G G G T T G T C C A  New ccp.seq15
751  . . C G G T G A A C G G G T T G T C C A  New cmvv34.seq5
1048 G G C G G G G A A C G G G T . G T C C A  New cmvwl.seq1
752  . . C G G T G A A C G G G T T G T C C A  New v27cp.seq5
736  . . C G G T G A A C G G G T T G T C C A  New v33cp.seq8

     T X X X X X X X X X X X X X X X X X X X  Majority
                       1170            1180
770  T . . . . . . . . . G G . . . . . . . .  carna5 cp cpexp33.seq
1142 T C C A G C T T A C G G C T A A A A T G  New ccp.seq15
769  T . . . . . . . . . . . . . . . . . . .  New cmvv34.seq5
1067 T C C A G C T T A C G G C T A A A A T G  New cmvwl.seq1
770  T . . . . . . . . . . . . . . . . . . .  New v27cp.seq5
771  T . . . . . . . . . . . . . . . . . . .  New v33cp.seq8

     X X X X X X X X X X X X X X X X X X X X  Majority
                       1190            1200
772  . . . . . . . . . . . . . . . . . . . .  carna5 cp cpexp33.seq
1162 G T C A . G T C G T G G A G A A A T C C  New ccp.seq15
770  . . . . . . . . . . . . . . . . . . . .  New cmvv34.seq5
1087 G T C G T G T C T T T C A . . . . . . C  New cmvwl.seq1
771  . . . . . . . . . . . . . . . . . . . .  New v27cp.seq5
772  . . . . . . . . . . . . . . . . . . . .  New v33cp.seq8
```

INTERNATIONAL SEARCH REPORT

International Application No
PC./US 95/07234

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/40 C12N15/82 C12N5/10 C07K14/08 A01H5/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C07K A01H

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	BIOTECHNOLOGY, vol. 10, December 1992 pages 1562-1570, GONSALVES, D., ET AL. 'COMPARISON OF COAT PROTEIN-MEDIATED AND GENETICALLY-DERIVED RESISTANCE IN CUCUMBERS TO INFECTION BY CUCUMBER MOSAIC VIRUS UNDER FIELD CONDITIONS WITH NATURAL CHALLENGE INOCULATIONS BY VECTORS' see the whole document --- -/--	1-36

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
23 October 1995	01-02-96

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+ 31-70) 340-3016	Maddox, A

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	PHYTOPATHOLOGY, vol. 81, no. 7, 1991 pages 794-802, QUEMADA, H.D., ET AL. 'EXPRESSIONOF COAT PROTEIN GENE FROM CUCUMBER MOSAIC VIRUDS STRAIN C IN TOBACCO: PROTECTION AGAINST INFECTIONS BY CMV STRAINS TRANSMITTED MECHANICALLY OR BY APHIDS' see the whole document ---	1-37
A	GENE , vol. 107, 1991 pages 181-188, NAMBA, S., ET AL. 'EXPRESSION OF THE GENE ENCODING THE COAT PROTEIN OF CUCUMBER MOSAIC VIRUS (CMV) STRAIN WL APPEARS TO PROVIDE PROTECTION TO TOBACCO PLANTS AGAINST INFECTION BY SEVERAL DIFFERENT CMV STRAINS' see the whole document ---	1-37
A	J. GEN. VIROL, vol. 74, 1993 pages 319-322, NAKAJIMA, M., ET AL 'PROTECTION AGAINST CUCUMBER MOSAIC VIRUS (CMV) STRAINS O AND Y AND CHRYSANTHEMUM MILD MOTTLE VIRUS IN TRANSGENIC TOBACCO PLANTS EXPRESSING CMV-O COAT PROTEIN' see the whole document ---	1-37
A	EP,A,0 412 912 (BIOSEM) 13 February 1991 see the whole document ---	1-37
A	EMBL ACC. NO. M98501. REL.33, 13-8-1992. CUCMBER MOSAIC VIRUS COAT PROTEIN (CP) RNA PAPPU, S.S., ET AL. see sequence ---	1-37
A	WO,A,89 05858 (UPJOHN CO) 29 June 1989 see claims 5-7 ---	10,11, 21,22, 32,33, 36,37
A	EP,A,0 480 310 (TAKEDA CHEMICAL INDUSTRIES) 15 April 1992 see the whole document ---	11,22, 33,37
A	WO,A,91 04332 (UPJOHN CO) 4 April 1991 see the whole document ---	10,21, 32,36

-/--

1

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	J. CELL. BIOCHEM. SUPPL., vol. 16F, 1992 page 222 TRICOLI, D. M., ET AL. 'TRANSGENIC SQUASH PLANTS EXHIBIT COAT PROTEIN MEDIATED PROTECTION UNDER FIELD CONDITIONS' see abstract Y232	38-49
A	J. CELL. BIOCHEM. SUPPL., vol. 18A, 1994 page 91 TRICOLI, D.M., ET AL. 'FIELD TRAIL RESULTS OF TRANSGENIC SQUASH AND CANTALOUPE PLANTS CONTAINING MULTIPLE VIRUS RESISTANCE' see abstract X1-126	38-49
A	WO,A,90 02184 (UPJOHN CO ;CORNELL RES FOUNDATION INC (US)) 8 March 1990 see the whole document	38-49

INTERNATIONAL SEARCH REPORT

ormation on patent family members

Interr 'onal Application No
PC1/US 95/07234

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
EP-A-0412912	13-02-91	FR-A-	2651504	08-03-91
		DE-D-	69007370	21-04-94
		DE-T-	69007370	06-10-94
		ES-T-	2063312	01-01-95
		JP-A-	3103127	30-04-91
		US-A-	5422259	06-06-95
WO-A-8905858	29-06-89	AT-T-	107361	15-07-94
		AU-B-	2927689	19-07-89
		DE-D-	3850261	21-07-94
		DE-T-	3850261	27-10-94
		EP-A-	0391972	17-10-90
		JP-T-	3501680	18-04-91
EP-A-480310	15-04-92	JP-A-	5192054	03-08-93
WO-A-9104332	04-04-91	AT-T-	114720	15-12-94
		AU-B-	6284090	18-04-91
		DE-D-	69014644	12-01-95
		DE-T-	69014644	04-05-95
		EP-A-	0491733	01-07-92
		ES-T-	2065545	16-02-95
		JP-T-	5500308	28-01-93
WO-A-9002184	08-03-90	AU-B-	639891	12-08-93
		AU-B-	3970489	23-03-90
		AU-B-	634168	18-02-93
		AU-B-	3987089	23-03-90
		CA-A-	1332718	25-10-94
		CA-A-	1329561	17-05-94
		DE-D-	68915282	16-06-94
		DE-T-	68915282	29-09-94
		EP-A-	0429478	05-06-91
		EP-A-	0429483	05-06-91
		JP-T-	4500151	16-01-92
		JP-T-	4500152	16-01-92
		WO-A-	9002189	08-03-90